

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2002, 00:20:11 ; Search time 3375.48 Seconds
(without alignments)
11345.220 Million cell updates/sec

Title: US-09-755-456-1

Perfect score: 1830
Sequence: 1 gaattcagatgcctcattacc.....tcccttagaagtgtggtac 1830

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenBank: *
1: gb_da: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_da: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_inv: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	1830	100.0	1830	6	AX007049	AX007049 Sequence
2	1830	100.0	1831	3	ECU5666	AJ005666 Encephal1
3	1828.4	99.9	220294	3	CNS07EGD	AL590446 Chromosom
4	1019	55.7	1019	3	AF310677	AF310677 Encephal1
5	1011.6	55.3	1018	3	AF310678	AF310678 Encephal1
6	844.8	46.2	936	3	AF310679	AF310679 Encephal1
7	176	9.6	1116	6	AX007051	AX007051 Sequence
8	144.2	7.9	1696	3	AF044915	AF044915 Encephal1
9	102.6	5.6	1340	3	AF338363	AF338363 Encephal1
10	101.4	5.5	1184	3	AF338361	AF338361 Encephal1
11	99.4	5.4	1297	3	AF338362	AF338362 Encephal1
12	99	5.4	1232	3	AY024342	AY024342 Encephal1
13	49.4	2.7	125020	9	AF429315	AF429315 Homo sapi
14	47.4	2.6	125020	9	AF429315	AF429315 Homo sapi
15	46.6	2.5	203108	9	CNS05TC3	AL355032 Human chr
16	46.6	2.5	207774	2	CNS0000H	AL049836 Human chr
17	46	2.5	66491	2	AC006715	AC006715 Caenorhab
18	46	2.5	89370	3	AC084153	AC084153 Caenorhab
19	44.4	2.4	686	8	CNS01DIR	AL116868 Botrytis
20	43.2	2.4	33499	9	AL590548	AL590548 Human DNA
21	43.2	2.4	163057	2	AC012419	AC012419 Homo sapi
22	43.2	2.4	163852	9	AC015971	AC015971 Homo sapi
23	43	2.3	3489	9	AF217688S2	AF217689 Homo sapi
24	43	2.3	257817	2	AC006909	AC006909 Caenorhab
25	42.4	2.3	211579	2	AL645845	AL645845 Mus muscu
26	42.2	2.3	170707	2	AL589986	AL589986 Homo sapi
27	42.2	2.3	178671	2	AL627110	AL627110 Homo sapi
28	42.2	2.3	199739	2	AC005528	AC005528 Mus muscu
29	41.4	2.3	37369	3	CEY26H2	282055 Caenorhabd1
30	41.4	2.3	56870	2	CEY39B6-3	Continuation (4 of
31	41.4	2.3	110000	2	CEY39B6-2	Continuation (3 of
32	40.6	2.2	7218	6	166494	166494 Sequence 14
33	40.4	2.2	34727	3	CEP11A6	281498 Caenorhabd1
34	39.6	2.2	67453	2	AC104900	AC104900 Mus muscu
35	39.6	2.2	172224	2	AC087149	AC087149 Mus muscu
36	39.6	2.2	190738	2	AC087222	AC087222 Homo sapi
37	39.6	2.2	205228	2	AC068584	AC068584 Homo sapi
38	39.4	2.2	75153	2	AC100376	AC100376 Mus muscu
39	39.4	2.2	103829	9	AP000665	AP000665 Homo sapi
40	39.4	2.2	160222	9	AP002962	AP002962 Homo sapi
41	39.2	2.1	1204	3	ASU249373	AJ249373 Ascaris s
42	39.2	2.1	175231	2	AC097084	AC097084 Rattus no
43	39	2.1	696	8	CNS01A92	AL112638 Botrytis
44	39	2.1	9673	2	AC104348	AC104348 Homo sapi
45	39	2.1	143291	9	HS163G9	AL008733 Human DNA

ALIGNMENTS

RESULT 1	AX007049	1830 bp	DNA	Linear	PAT 06-SEP-2000
LOCUS	AX007049	Sequence 1 from Patent WO0001724.			
DEFINITION	AX007049				
ACCESSION	AX007049.1	GI:994975			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM	Encephalitozoon cuniculi.				
REFERENCE	Encephalitozoon cuniculi				
AUTHORS	Eukaryota: Microsporidia; Unikaryonidae; Encephalitozoon.				
TITLE	1 (bases 1 to 1830)				
JOURNAL	Vivares, C., Danchin, A. and Delbac, F.				
FEATURES	Microsporidium polar tube proteins, nucleic acids coding for said				
source	Proteins and their uses				
BASE COUNT	Patent: WO 0001724-A 1 13-JAN-2000;				
ORIGIN	VIVARES CHRISTIAN (FR); DANCHIN ANTOINE (FR); DELBAC FREDERIC (FR);				
	CENTRE NAT RECH SCIENT (FR)				
	Location/Qualifiers				
	1..1830				
	/organism="Encephalitozoon cuniculi"				
	/db_xref="taxon:6035"				
	460 a 486 c 434 g 450 t				

Query Match 100.0%; Score 1830; DB 6; Length 1830;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1830; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gaataagatgctcatcacttggagattaaaaatgattgcttattgtatatactcg 60
 |||||||
 Db 1 GAATTAGATGCTCATACCTTGGATTTAAAAATTGATTCATTGTATATATCCTG 60

QY 61 ggcggagacggccggcctgcttcttcagggggtgtccctaacccagatgcacaggaatttc 120
 |||||||
 Db 61 GGGGAGACGGCCGGCTGCTGATTTCTCAGGGGTGTCCCTTACCCAGTGACAGGAGGTTC 120

QY 121 ggaagtgcttggatgaaagtaagccatttgggttctcatccatgcatcgtccct 180
 |||||||
 Db 121 GGAAGTGTCTGTGATGGAAGTAAGGCCATTTGTGGTTCTCATGTCATGTCATGCCCT 180

QY 181 ttcggtctttccaagaatccaattatctcccaaggacttcaaccctcagaatggaa 240
 |||||||
 Db 181 TTCGGCTGTTCCACAAAGATCCAAATTATTCCTCAGGACTTTCACACCTTCAGAAATGGAAA 240

QY 241 cagaagatgaactctctgtgcaaatctgataatgataatgagagatgaatgaaccagag 300
 |||||||
 Db 241 CAGAGATGAACCTCTGTGCAATGCTGATATGCAATTTGGAGACATGTGAACCCAGGAG 300

QY 301 ttgtaataaagtaataatactcccgaaacgcagagtttaagatgaagatttctta 360
 |||||||
 Db 301 TTGGAATAAAGTATATAATACCTCCGAAAGCAGAGTTTAAAGTAAGAAAGTATTTCTA 360

QY 361 agatcctctctgctctatctgacctgtgaatgtggaagatgtctatccaagacgcgc 420
 |||||||
 Db 361 AGATCCTCTCTGCTCTATTTGGCCCTGATGAAGTTGGAATGCTATTCAGCAACCGCAC 420

QY 421 tgtgcagcaatgataatgacctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 480
 |||||||
 Db 421 TGTGACGACATGATATGCTGCTTACTCCGGGACACAGGATGAGCTCAGCACCCCTGCT 480

QY 481 atgtgctgatacccaagaccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 540
 |||||||
 Db 481 ATGTGCTGATCCCGACACCCCGGGAACCATAGCAAACTGTGCACGCGTTTACAGAGACA 540

QY 541 catattctctctctccgcgtgcacccacatctccagtgactccgggggaaactacagaga 600
 |||||||
 Db 541 CATATTCTCTTCTCCCGCTGACCCACCATCTCAGTGAATCCGGGGAACCACTAGGAGA 600

QY 601 atagaacatctccatcgctcgtcgcagaagatgtagaacaatgcaagattgcgtattga 660
 |||||||
 Db 601 ATAGACATCTCCATGCGCTCCCTGCAGAAAGATGTAGGAACATGCAAGATTGCCGTATTGA 660

QY 661 agcactgcagacgcacaggaacaacalcagggagacacacagggcttctgtga 720
 |||||||
 Db 661 AGCAGTCGCGACGACAGGACGAACATCAGGAGACACACAGGGGTCTGTGAAA 720

QY 721 ccccaagacagcaagccttctgtaagtgatctccacacacacacacacacacacacac 780
 |||||||
 Db 721 CCCCAAGACGACAGACCTTTGTGAGTATCTCCACCACTCTCGCGTACCGGTGACTG 780

QY 781 tggagctgcagacagcttccatctgtgtgcaagttgttccgtgctgtgctcaccacag 840
 |||||||
 Db 781 TGGAGCTGTGACAGCTCCATCTGTTGTGTGCACTGTGCTCTGCTGCTCACCACACAG 840

QY 841 cagttccaggctactacaacaatgaaacalcggtatctccgagacagcaacagatcttt 900
 |||||||
 Db 841 CAGTTCAGGCTACTACAAATGAACATCCGGTATCTCTGACAGCAACAGATCTTT 900

QY 901 ctggacactctctcccaagagcacttctgtgtgaaggaacagccatgctactgactcct 960
 |||||||
 Db 901 CTGGCACTCTTCCCAAGGACCACTTGTGTGACAGGACAGGCAATGCCATGACCTCTG 960

QY 961 gacagcaacagatcctcttgcgactcttcccccagggggtcacttctgtctcagggagag 1020
 |||||||
 Db 961 GACAGCAACAGATCTTTCTGCGACTCTTCCCCAGGGGTCACTTTGTGTGACGGACAG 1020

QY 1021 ccagcctagcactcctctgggcagcaacaggtccttcttggcactcttccccaggaagta 1080
 |||||||
 Db 1021 CCACGCGTATGCACTCTCTGGGAGCAACAGTCTCTTCTGGCACTCTCTCCCCAGGAGTCA 1080

QY 1081 cttgtgtcaaggagcagccacgcctagcactcctctgggcagcaacaggtccttctgga 1140
 |||||||
 Db 1081 CTTTGTGTGAGGAGCAGGCGCAGCCCTAGCACTCTGGGAGCAACAGGTCCTTTCTGGGA 1140

QY 1141 cctctcccaaggagcacttctgttcaagatcaagatgtgctgtaaacatccgagttc 1200
 |||||||
 Db 1141 CCTTCTCCAGAGGACCACTTTGTGTGAGATCAAGTATGCTCTGAGACATCCGAGTTTC 1200

QY 1201 ctggacagcagggacagcttctgtagcagtgctgtgtgcccctcagaattccaacctgtca 1260
 |||||||
 Db 1201 CTGGACAGCAGGAGCAGTCTAGTGACAGTGTGTGTGCCCTCAGATTCACAAACCTGTCA 1260

QY 1261 tgcgcacatccatgaacatgaatggaatggtatccctctctacgcatacagccaa 1320
 |||||||
 Db 1261 TGCCGCACTCCATGAACATTTAGTGAATGGGTATCTTCTTACCCGATACAGCCCA 1320

QY 1321 acctcgatcactggagatctgtttagacatacagaagaacggggagacatctgcgagc 1380
 |||||||
 Db 1321 ACCTCGATCACTGGGATCTGTGTGACATACAGAACGGGGAGCATCTGCGAGC 1380

QY 1381 aaaaaccgagaagctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1440
 |||||||
 Db 1381 AAAAAACCGAGAAGTCCGCGACGACGATGCCATGAGGCTGTGCAACACCAACCA 1440

QY 1441 cggttatataaggcaacagcaggtatctctgttggaccaggaatgtacaatgaatgaat 1500
 |||||||
 Db 1441 CGGTATTTATAGCAACACGAGTATCTTGTGAGCAGAGAAATGTAATGCAATTAATCT 1500

QY 1501 ctccatgaacacactgtctgcacatgtctgtagcctaataaacaagattatcttct 1560
 |||||||
 Db 1501 CTCCATGCAACACTGCTGTGCAATGCTGCTAGCTTAATAAACAAGTATTAATCTTCT 1560

QY 1561 tttctcggctcttcttggaaagtgtggaatggaagagctatgagctgaagtgaaat 1620
 |||||||
 Db 1561 TTTCTCTGCTCTTTGGAAGCTTGGATGGGATGAGAGATCTATGAGCTGAAGTGAAT 1620

QY 1621 gccaacactctctctgcccagaacacatctgagatgttctctctgtgccaagagatttg 1680
 |||||||
 Db 1621 GCCAACACTCTCTCTGCCAAGAACATTCGATGTTCTCTGCGCAGAGATTTTG 1680

QY 1681 taacagatcccccagagatttagcagccttggagatcatgatgaatcaatgaatgaat 1740
 |||||||
 Db 1681 TAACAGATTTCCCGAGATTTTAGAGAGCTTTGAGATGATGATGAATCAAGTATTAAC 1740

QY 1741 ttctcaattatttattcttctgttataatcccgagccaactgaaagaatgctc 1800
 |||||||
 Db 1741 TTCTCAATTATTATTATCTTCTGTTTATATCCGAGCAATCTGAGAGAAATGCTCTC 1800

QY 1801 gaattcaagctcccttagaagatgtggatc 1830
 |||||||
 Db 1801 GAATTCAAGCTCCCTTAGAAGTGTGGATC 1830

RESULT 2
 ECUS666 1831 bp DNA linear INV 04-SEP-1998
 LOCUS ECUS666
 DEFINITION Encephalitozoon cuniculi gene encoding polar tube protein.
 ACCESSION AJ005666
 VERSION AJ005666.1 GI:3451140
 KEYWORDS polar tube protein.
 SOURCE Encephalitozoon cuniculi.
 ORGANISM Encephalitozoon cuniculi
 Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.
 REFERENCE 1 (bases 1 to 1831)
 AUTHORS Delbac F.
 TITLE Direct Submission
 JOURNAL Submitted (21-APR-1998) Delbac F., Universite Blaise Pascal UPESA
 CNRS 6023, Laboratoire de Biologie Comptee des Protistes, 24

avenue des Landais, 63177 Aubiere, FRANCE
2 (bases 1 to 1831)
Delbac, F., Peyret, P., Metenier, G., David, D., Danchin, A. and
Vivares, C.P.
On proteins of the microsporidian invasive apparatus: complete
sequence of a polar tube protein of *Encephalitozoon cuniculi*
JOURNAL
MOL. Microbiol. 29 (3), 825-834 (1998)
98389314
FEATURES
source
1. .1831
/organism="Encephalitozoon cuniculi"
/db_xref="taxon:6035"
345. .410
345. .1532
/codon_start=-1
/product="polar tube protein"
/protein_id="CA006662.1"
/db_xref="GI:3451141"
/db_xref="SPTREMBL:076942"
/translation="MKGISKILISASIALMKLENYSATALCSNAYGLTPQOGMAQOP
SYVLIPSTPPTIANCASGSQDTYSPSPAATSPVTPGKTSENETSPAPADYGTCKI
AVLKHCDAPCTTSGTTPGSGPCEPPEOQPLSYSTPAPVTVESQSPVVPVY
VAHNOAVPGYNNGTSGIPGQOOLISGLTPPGATLCOGMAPSPGQOOLISGLTPG
VTCQQAATPSTPGQOQVLSGTIPRPTVLCQGAATPSTPGQOQVLSGLTPGATLCD
QGMPTSGVPGQOQSGQSCAPQIPMPVPSMNTISNGTSPSTAYSPMLGSLGSCV
DIQGTGTSCEQRPERSAIQYAMEACATPTPIYIIGNSEYLVPGMNAINSPONTAV
QCC"
mat_peptide
411. .1529
BASE COUNT 460 a 487 c 434 g 450 t
ORIGIN
Query Match 100.0%; Score 1830; DB 3; Length 1831;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1830; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 gaattcagatgctcattaccttggatgataaaatgtagtgcatttggatataatcctg 60
Db 1 GAATTGAGATGCCATACCTTGGGATTAAAAAATTGATGTTCAATTGTATATATCTG 60
QY 61 ggagagacagcgccgctcgtattcttcaggagggtgcgcctaccacgtgacaggggtcc 120
Db 61 GGCAGACAGCGCGGCTGATCTTTCAGGGGTGTCCTACCCAGTGACAGAGGTTCC 120
QY 121 ggaaggtctctgtagtggaagtaagggcatttgggtgttcacatcattcgtccct 180
Db 121 GGAAGGTCTCTGATGGAAGTAAGGCCATTGTGGGTTCTCATCCATGTCATGCTCCT 180
QY 181 ttgggtcttaccagaatcgaattatctcctcaggaattcgaacccctcagaatgaaa 240
Db 181 TTGCGCTCTTCCACCAAGATCCAAATTATCTCTCCAGCACTTCAACCTCAGAAATGAAA 240
QY 241 cagagatgaaactctctgtgcaaatcgtagaatcagatggaagacattgaaacacagag 300
Db 241 CAGAGATGAACCTCTGTGTGAATACTGAGATATGATGAGACATTGAAACACCGAG 300
QY 301 ttctgaaataaagataaatacctccgaaaaacgagattcaagaatgaagtaattctta 360
Db 301 TTTGAAATTAAGATATAATACCTCCGAAAACGAGATTTAAGATGAAGATATTCTTA 360
QY 361 agatcctctctgctctatttgcctgtagaagtggagaatgctcattcaagcaacggcac 420
Db 361 AGATCCTCTCTGCTCTATTGCTCTGATGAGAGTGAAGAAATGTCATTACACCAACCGCAC 420
QY 421 ttgtcagaatgcatatgagcctaactccggagacaacagggatgagctcagaacggctgt 480
Db 421 TGTCAAGCAATGCAATAGGCTTAATCTCGGAGAACAGGATATGGCTCAGACGCCGTGT 480
QY 481 atgtctgataccacagaccccggaacacatagcaaacctgtgcaagcggttccacagaca 540
Db 481 ATGTGTGATCCCAAGACCCCGGGAACCAATAGCAAAATGTGCAAGCGGTTCCACAGACA 540

QY 541 catattcctcttccgcctgcacaccacatctccagtgacttcggyggaatactagcgaga 600
Db 541 CATATTCTCTTCCCTCCGCTGACACCATCTCCAGTACTCCGGGAAAACTAGCGAGA 600
QY 601 atgagacttctccatctgctctcctcgagaagatgtatgaacttgcagaatttcctatta 660
Db 601 ATGAGACTCTCTCATCGGCTCTCGAAGAAATGTAAGAACATGCAAAATTCCTGATTGA 660
QY 661 agcactgcagcagcaacaaacaaatcaggagacagacacaggggtcaggccttctgaaa 720
Db 661 AGCACTGGAGCAGCACACAGAAACAAATCAGGAGACACACAGGCTCAGGCTTGTGAAA 720
QY 721 cccagagacagcaaacagccttctgtaagtatctccaccactctgcgtacagctgactg 780
Db 721 CCCAGAGACAGCAACAGCCTTGTGATGATCTCCACCACTCTCTCCGCTGACTGTG 780
QY 781 tggagctcgcagacttccatctctgttggccaagtgttccctgtgttctcaaccacgg 840
Db 781 TGGAGTCTGCACAGTCTCATCTGTTGTGCCAGTGTTCGTGCTCTCACCCACAG 840
QY 841 caattccagctactacacaatggaacatccggtattccttggacagacagactcctt 900
Db 841 CAGTTCCAGGCTACTACACAATGGAACATCCGGATTCTTGACAGACAGATCTTT 900
QY 901 ctgagactcttcccccaggaagccaatttgtlcaaggagacagccatgcttagcactctg 960
Db 901 CTGACACTCTTCCCCAGAGCACTTGTGTACAGGACAGGCCATGCTAGCACTCTG 960
QY 961 gacagcaacaaatccttctctgcaactctcccccaggggttaacttgtgtcaaggacag 1020
Db 961 GACAGCAACAAATCTTCTTGACACTCTCCCCAGGGGTCACTTGTGTGACAGGACAG 1020
QY 1021 ccaagcctagcaactccttggagacagacaggtccttcttgcgacacttccccaagagatca 1080
Db 1021 CCAGCGCTAGCACTCTCTGGGACAGACAGGTCTTTTCTGCGACTCTTCCCGAGAGTCA 1080
QY 1081 ctgtgtlcaaggagacagccagccttagcactccttggcagaacaggttcttcttgca 1140
Db 1081 CTTGTGTCAAGGAGACAGCCACGCTTACACACTCTGGGACAGAACAGGCTTCTTGCA 1140
QY 1141 ccccttccagggagccacttgtgtcagaatcgaagtatgctgtgaacatcccgagttc 1200
Db 1141 CCTTCTCCAGGAGCCACTTGTGTGACAGATCAAGATGATGGAACATCTCCGAGATC 1200
QY 1201 ctggaacagaggaacagcttagtgagacagtgttggcccctcaagaattcacaacccctgtca 1260
Db 1201 CTGGAACAGGAGACAGCTTAGTGACAGTGTGCCCCCTCAGATTCCAAAACCTGTGCA 1260
QY 1261 tgcgcacatcatgaaacattagtggaaatgggtatccttctctctacacgcatacagcccaa 1320
Db 1261 TGCCGCCATTCATGAACAATTAGTGAATGGGTATCTTCTTACCCGATACAGCCCAA 1320
QY 1321 aacctggtactgggtacccgtgtgtgtgacatacagaagaacgggggtgacatccctcgagc 1380
Db 1321 AACTGTGATCACTGGGATCTGTGTGACATACAAAGACGGGGGATCACTCTCGAGC 1380
QY 1381 aaaaacccgaagaagtcgacagcagtagtgcataaggagcctgtgcaacacaacacaa 1440
Db 1381 AAAAAACCGAAGAGTCCGCCACAGCAAGTATGCCATGAGAGCCCTGTCAACACCAACCA 1440
QY 1441 cggttattataggcaacagcagtagtcttgttggacccaaggaatgtacaaatgaat 1500
Db 1441 CGGTATTATATAGGCAACGCGAGTATCTTGTGACACAGGAATGTAACGAATTAATCT 1500
QY 1501 ctccatgcaaacactgcttccaatgtctgtggtcctaataaagaattcaatctctt 1560
Db 1501 CTCATGCAACACTGCTCTCCAAATGCTCTAGGCTTAATAAAGAAATTAATCTTCTT 1560
QY 1561 ttctctcgctcttggagagcttggatgggaatggagagctatagggctgaagtgaat 1620
Db 1561 TTTCTTGGCTCTTTTGGAAAGCTTGAGATGGGATGAGAGAGTCTAATGGCTGAAGTAAT 1620
QY 1621 gccaaactcttctgcccagaacacatcggatgttcttctctgttggcaggagtttg 1680

|||||
Db 1621 GCCAAGCTTCTTCCGAGAACACATTCGGATGTCCTGTCGGCCAGGATTTGG 1680
Oy 1681 taacagattccccgagattttagcagccttggaglacatgattgaatcagattaaac 1740
|||||
Db 1681 TAACAGAGATTCCCCGAGATTAGCAGCCTTGAGATACCATGATGATCAGATTAAAC 1740
Oy 1741 ttctcaattatttattcttcttcttattatcccgagccaactgagaagaatgcttc 1800
|||||
Db 1741 TTCTCAAAATTAATTAATTTCTTCTGTTTATATCCGAGCAATCGAAGAAATGCTTC 1800
Oy 1801 gaattcaagctccctagaagtgatgac 1830
|||||
Db 1801 GAATCAAGCTCCCTTAAGATGTTGGATC 1830
RESULT 3
LOCUS CNS07EGD 220294 bp DNA linear INV 23-NOV-2001
DEFINITION chromosome VI of strain GB-M1 of Encephalitozoon cuniculi
(Microspore).
ACCESSION AL590446
VERSION AL590446.1 GI:13560073
KEYWORDS Chromosome sequence.
SOURCE Encephalitozoon cuniculi.
ORGANISM Encephalitozoon cuniculi.
REFERENCE 1 (bases 1 to 220294)
Peyret, P., Katinka, M.D., Duprat, S., Duffieux, F., Barbe, V.,
Barbazanges, M., Weissenbach, J., Saurin, W. and Vivares, C.P.
AUTHORS Direct Submission
TITLE Submitted (05-APR-2001) to the EMBL/GenBank/DBJ databases
REFERENCE 2 (bases 1 to 220294)
Genoscope.
JOURNAL Submitted (05-APR-2001) Genoscope - Centre National de Sequencage,
2 rue Gaston Cremieux, CP 5706, 91057 Evry cedex, FRANCE. (E-mail :
sequref@genoscope.cns.fr Web : www.genoscope.cns.fr)
EQUIPE de Parasitologie Molculaire et Cellulaire, UPRES A CNRS
6023, Universite Blaise Pascal, 63177 Aubiere cedex, France (E-mail
: Pierre.PEYRET@pdp.univ-bpclermont.fr; Tel : (33) 4.73.40.74.34;
Fax : (33) 4.73.40.76.70).
FEATURES
Location/Qualifiers
source 1..220294
/organism="Encephalitozoon cuniculi"
/strain="GB-M1"
/db_xref="taxon:6035"
/chromosome="VI"
BASE COUNT 57653 a 52579 c 51594 g 58468 t
ORIGIN
Query Match 99.9% Score 1828.4; DB 3; Length 220294;
Best Local Similarity 99.9% Pred. No. 0;
Matches 1829; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 1 gaattcaagctccctagaagtgatgac 1830
|||||
Db 36787 GAATTCAGATGCTCATACCTTGGATTAATAAATGATGTCATTTGTTATATATCCCTG 36728
Oy 61 ggggagcaagcgagctcgatattcttcaaggggtgcgctcctaccagtgcaagaaggttcc 120
|||||
Db 36727 GGGGAGCAAGCGCGCTCGATTTCTCAGGGGTGTCGCTACCCAGTGCAAGAGAGGTTC 36668
Oy 121 ggaagtgctcttgatggaagaagcatttgggttcctacatgcatgcgtccct 180
|||||
Db 36667 GGAAGTGCTCTTGATGGAAGTAAGCCATTTTGGGTTCTCATGTCATGTCGTCCT 36608
Oy 181 ttgcgctgtttcccaagatccaatattcttcccaagacttccaaacctcagaatgga 240
|||||
Db 36607 TTCAGCTGTTTCACCAAGATCCAAATATTCCTCCAGAGACTTCAACCCCTCAAGATGGA 36548
Oy 241 cagagatgaactctctctgtgcaaatcgtagatatcgattggagaattgaaccaagga 300

|||||
Db 36547 CAGAGATGAAGAACTCTGTGCAATGATATCATTTGGAACATTTAAACACAGGGAG 36488
Oy 301 ttggaataaagfataaataacctccgaagaacgagaatttgaagatgaaggtattcta 360
|||||
Db 36487 TTTGAATATAAAGTATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 36428
Oy 361 agatctctctgctctatcttcttcttcttcttcttcttcttcttcttcttcttct 420
|||||
Db 36427 AGATCTCTCTGCTCTCTATGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 36368
Oy 421 ttgtgcagcaatgataatgctccttcttcttcttcttcttcttcttcttcttcttct 480
|||||
Db 36367 TGTGACAGCAATGATATGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 36308
Oy 481 atgtgcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 540
|||||
Db 36307 ATGTGCTGATCCCCACACACACACACACACACACACACACACACACACACACAC 36248
Oy 541 catattctcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 600
|||||
Db 36247 CATATTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 36188
Oy 601 atgagacatctcctatcgctctcttcttcttcttcttcttcttcttcttcttcttct 660
|||||
Db 36187 ATGAGACATCTCCATCGGCTCTGAGAGATGTAAGATGCAAGATGCAAGATGCAAG 36128
Oy 661 agcactgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 720
|||||
Db 36127 AGCACTGCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 36068
Oy 721 ccccaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 780
|||||
Db 36067 CCCCAAGAGCAGCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 36008
Oy 781 tgaagctgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 840
|||||
Db 36007 TGAAGTCTGACAGTCTGACAGTCTGACAGTCTGACAGTCTGACAGTCTGACAG 35948
Oy 841 caattcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 900
|||||
Db 35947 CAGTTTCAGAGCTCTGACAGCTCTGACAGCTCTGACAGCTCTGACAGCTCTGACAG 35888
Oy 901 ctggcactcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 960
|||||
Db 35887 CTGGCACTCTTCCCGAGGAGGACCTTTGTCAGGAGACAGGACCTTCTGACACTCTG 35828
Oy 961 gacagcaaacagatcttcttcttcttcttcttcttcttcttcttcttcttcttctt 1020
|||||
Db 35827 GACAGCAACAGATCTTCTGACAGCTCTGACAGCTCTGACAGCTCTGACAGCTCTG 35768
Oy 1021 ccaagcctagcactcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1080
|||||
Db 35767 CCACGCTTAGACATCTCTGAGGAGCAACAGGTCCTTTCTGACACTCTTCCCGAGG 35708
Oy 1081 ctttgtcagaggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1140
|||||
Db 35707 CTTTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 35648
Oy 1141 ccttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1200
|||||
Db 35647 CCTTCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 35588
Oy 1201 ctggacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1260
|||||
Db 35587 CTGCAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 35528
Oy 1261 ttgcgcacatccatgaacattagtggaaatggatcttcttcttcttcttcttcttct 1320
|||||
Db 35527 TGCAGCATTCATGAACATTAAGTGAATAGGATACCTTCTTCTACCGCATACAGCCCA 35468
Oy 1321 acctggatcactggatctctgtgtgataaagaagcgggggggagacatctcgagac 1380
|||||

D3	35467	ACCTCGGATTCACGTGGGATCCCTGTGTGACATACAGAAACCGGGGGGACATCTCGCAGC	35408				
OY	1391	aaaaacccgaagaagtcgcgcgcacgcagta tgcgcataagagcctgtgcaacaccacaacaa	1440				
D3	35407	AAAAACCCGAGAAAGTCCGCCACGAGTATGCGCATGAGAGCGCTGTGCAACACCAACACCA	35348				
OY	1441	cgggtataatagggcaacagcgagatctctgtgtgacaggaatgtacaa tgcataaact	1500				
D3	35347	CGGTATATATAGGCAACAGCGAGTATCTTGTGTGGACCGAGATGTACAAATGCAATTAAC	35288				
OY	1501	ctccatgcaaacacgcctctccaa tgcctcctcctggcctaaataaaacagat ttaactctt	1560				
D3	35287	CTCCATGCAACACGCTCTCCAAATGCTCTAGCGCTAAATAATTAACGAGCTTTAACTCTCT	35228				
OY	1561	ttctctcgctctcttggaacgttgatgtaggagatgagagagctatgagctgaagtgaat	1620				
D3	35227	TTTCTCGGCTTTTGGAAACGTGGATGGGATGAGAGAGTCTATGGGCTGAAGTGAAT	35168				
OY	1621	gccaacactctctctgccaagaacacac tggagatgtctctctctgtgccaagatttg	1680				
D3	35167	GCCAACTCTCTCTGCCCAAGAACATTCGAGTGTCTTCCTGTGCGCCAGAGATTGG	35108				
OY	1681	taacaggatctcccgaggagattagagagctgtagtaccatgatgatgaatcgtatataac	1740				
D3	35107	TAAACGAGATCCCGCAGATTTTACGACGCTTGAGTACCATGATTAATGATTAATTAAC	35048				
OY	1741	ttctcaaatattatctctctctgttcttatacccggaacatctgagaagatgctc	1800				
D3	35047	TTCTCAATATATATATCTTCTGTCTTTATATATCCCGACCAATCTGAAAGATGCTC	34988				
OY	1801	gaattcaagctcccttagaagtgtyggatc	1830				
D3	34987	GAAATCAAGCTCCCTTAAAGATGTGGATC	34958				
RESULT 4	AF310677	Encephalitozoon cuniculi I polar tube protein gene, partial cds.					
LOCUS	AF310677	Encephalitozoon cuniculi I polar tube protein gene, partial cds.	1019 bp	DNA	linear	INV 17-JUN-2001	
DEFINITION	AF310677	Encephalitozoon cuniculi I polar tube protein gene, partial cds.					
ACCESSION	AF310677	Encephalitozoon cuniculi I polar tube protein gene, partial cds.					
VERSION	AF310677.1	GI:14484923					
KEYWORDS							
SOURCE		Encephalitozoon cuniculi.					
ORGANISM		Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.					
REFERENCE		1 (bases 1 to 1019)					
AUTHORS		Xiao, L., Li, L., Vivasvaya, G. S., Moura, H., Didier, E. S. and Lal, A. A.					
TITLE		Genotyping Encephalitozoon cuniculi by Multilocus Analyses of Genes					
JOURNAL		With Repetitive Sequences					
PUBMED		J. Clin. Microbiol. 39 (6), 2248-2253 (2001)					
REFERENCE		2 (bases 1 to 1019)					
AUTHORS		Xiao, L., Li, L., Vivasvaya, G. S., Moura, H. and Lal, A. A.					
TITLE		Direct Submission					
JOURNAL		Submitted (04-OCT-2000) Division of Parasitic Diseases, National					
PUBMED		Center for Infectious Diseases, Centers for Disease Control and					
REFERENCE		Prevention, 4770 Buford Highway, Atlanta, GA 30341, USA					
FEATURES		Location/Qualifiers					
SOURCE		1..1019					
		/organism="Encephalitozoon cuniculi"					
		/isolate="CDC-V385"					
		/specific_host="Homo sapiens"					
		/db_xref="taxon:6035"					
		/note="genotype: I"					
		<1..>1019					
		/product="polar tube protein"					
		<1..>1019					
		/codon_start=3					
		/product="polar tube protein"					
		/protein_id="AAK62814.1"					
		/db_xref="GI:14484924"					
		/translation="SIALMKLENYYSATLCSNAYGLTGPCQCGAAGPQSVYLIPSTFG					
CDS							

Query Match	Best Local Similarity	Score 1019;	DB 3;	Length 1019;
Matches 1019;	Conservative	0;	Mismatches	0;
Indels	0;	Gaps	0;	
QY 373	cctctatgtcccttgaatgtgaatgtgaatgtctatctatgcagaacccgactgtgcagcaatg	432		
DB 1	cctctatgtcccttgaatgtgaatgtgaatgtctatctatgcagaacccgactgtgcagcaatg	60		
QY 433	catatgtgccttaactccgggacacacagggtaatgtctcaagcagccgtctcgtatgtctatcc	492		
DB 61	catatgtgccttaactccgggacacacagggtaatgtctcaagcagccgtctcgtatgtctatcc	120		
QY 493	ccagacaccccgaggaaacctatgaacactgttgcgaagcggttcacaggaacatactctctc	552		
DB 121	ccagacaccccgaggaaacctatgaacactgttgcgaagcggttcacaggaacatactctctc	180		
QY 553	ctcccgctgtcacccacatctctcagtgactccgggaaacacagcgaagaataagacatctc	612		
DB 181	ctcccgctgtcacccacatctctcagtgactccgggaaacacagcgaagaataagacatctc	240		
QY 613	catcgctctctgtcgaagaagatgtagaagaatgtcgaagatgtccgtatctgaagcactgcagc	672		
DB 241	catcgctctctgtcgaagaagatgtagaagaatgtcgaagatgtccgtatctgaagcactgcagc	300		
QY 673	caccaagaacaacatcgaaggagacacccaaggttcaggccttgtgaaccaccaagacagc	732		
DB 301	caccaagaacaacatcgaaggagacacccaaggttcaggccttgtgtgaaaccccagacagc	360		
QY 733	aacagaccttgtcagtgatctccacacactctccgcgcacccggttgactgtgaagtcagac	792		
DB 361	aacagaccttgtcagtgatctccacacactctccgcgcacccggttgactgtgaagtcagac	420		
QY 793	agtcctcatctgtgtgcagagtgtctctgtctgtctgtctgtctgtctgtctgtctgtctgt	852		
DB 421	agtcctcatctgtgtgcagagtgtctctgtctgtctgtctgtctgtctgtctgtctgtctgt	480		
QY 853	actacaacaatgtgaacatccggtatctctgtgacagcaacagatctcttctgtgacatctc	912		
DB 481	actacaacaatgtgaacatccggtatctctgtgacagcaacagatctcttctgtgacatctc	540		
QY 913	ccccagagagccacttgtgtctcaaggagacatgtcctcagacatctctgtgacagcaacaga	972		
DB 541	ccccagagagccacttgtgtctcaaggagacatgtcctcagacatctctgtgacagcaacaga	600		
QY 973	tcccttctgtgacatcttcccccaggagatcacttgtgtcagggagacaggcacgcctctaga	1032		
DB 601	tcccttctgtgacatcttcccccaggagatcacttgtgtcagggagacaggcacgcctctaga	660		
QY 1033	ctcctctgtgcagaaacaggttcttcttctgtgacatcttcccccaggagatcacttgtgtcagg	1092		
DB 661	ctcctctgtgcagaaacaggttcttcttctgtgacatcttcccccaggagatcacttgtgtcagg	720		
QY 1093	gaacagagccaagcctctagcactctctgtgacagcaacaggtccttcttctgtgacacccctctccag	1152		
DB 721	gaacagagccaagcctctagcactctctgtgacagcaacaggtccttcttctgtgacacccctctccag	780		
QY 1153	gaagcacttgtgtcagaagatcaaggtatgctcctgtgacaaatctcggagttcttctgtgacagcagg	1212		
DB 781	gaagcacttgtgtcagaagatcaaggtatgctcctgtgacaaatctcggagttcttctgtgacagcagg	840		
QY 1213	gaacagtcctagtggaacagtgttgtgtccctctcagatctccaaacccctgtctatgtccgcacatca	1272		
DB 841	gaacagtcctagtggaacagtgttgtgtccctctcagatctccaaacccctgtctatgtccgcacatca	900		

QY 1273 tgaacattagtggaatgggtatcctctctacccgatacagcccaaacctcgatcac 1332
|||||
Db 901 TGAACATTAGTGAATGGGTATCTCTCTACCCGATACAGCCCAAACTCGGATCAC 960
QY 1333 tgggataccgtgtgtgacatacagaagaacggggggagacatctcgcgagcaaaacccgag 1331
|||||
Db 961 TGGGATCTCTGTGTGACATACAGAAAGCGGGGAGACATCTCGAGCAAAAACCCGAG 1019

RESULT 5
AF310678 1018 bp DNA linear INV 17-JUN-2001
LOCUS Encephalitozoon cuniculi II polar tube protein gene, partial cds.
DEFINITION AF310678
ACCESSION AF310678.1 GI:14484925
VERSION
KEYWORDS
SOURCE Encephalitozoon cuniculi.
ORGANISM Encephalitozoon cuniculi
Eukaryota; Microsporidia; Unikaryoniidae; Encephalitozoon.
REFERENCE 1 (bases 1 to 1018)
AUTHORS Xiao,L., Li,L., Visvesvara,G.S., Moura,H., Didier,E.S. and Lal,A.A.
TITLE Genotyping Encephalitozoon cuniculi by Multilocus Analyses of Genes
with Repetitive Sequences
J. Clin. Microbiol. 39 (6), 2248-2253 (2001)
JOURNAL
MEDLINE 21270266
PUBMED 11376065
REFERENCE 2 (bases 1 to 1018)
AUTHORS Xiao,L., Li,L., Visvesvara,G.S., Moura,H. and Lal,A.A.
TITLE Direct Submission
Submitted (04-OCT-2000) Division of Parasitic Diseases, National
Center for Infectious Diseases, Centers for Disease Control and
Prevention, 4770 Buford Highway, Atlanta, GA 30341, USA
FEATURES
source
1..1018
Location/Qualifiers
1..1018
/organism="Encephalitozoon cuniculi"
/isolate="m104"
/specific_host="mouse"
/db_xref="taxon:6035"
/note="genotype: II"
<1..>1018
/product="polar tube protein"
<1..>1018
/codon_start=-3
/product="polar tube protein"
/protein_id="AAK62815.1"
/db_xref="GI:14484925"
/translation="TALMKLENYSATLCSNAYGTPGQGMADPOPSVYLIPSPGT
IANCASGSDPTSPSPAPPTSPYTBKTSSENTSPAPREDYGTCKIAYLAKCDABGT
TSGTTPGSPPTPEQOQPLSYISTTPAVTVESKQSPVVPVVAHVHDAVPQY
NNGTSGIPGQOQILSGTLPFGATLCQGMAMPSTPGQOQVLSGTLPGVYLICGQAPTS
TPGQOQVLSGTLPGVTLICGQAPSTPGQOQVLSGTLPGATLCCODDMPSTGVP
OQSGSGCCAPQIPNPVMPSPMISGNYPSSTAYSPMLGSLGSCVDIQKRGTSCE
OKPE"

BASE COUNT 237 a 309 c 254 g 218 t
ORIGIN

Query Match 55.3%; Score 1011.6; DB 3; Length 1018;
Best local Similarity 99.6%; Pred. No. 0;
Matches 1014; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 376 ctatgcacctgataagttggaatgctctatcacaacacgactgtgcaaatgcat 435
|||||
Db 1 CTAATGCCCTGATGAGATTGAGAAATGCTCATTCAGGAACCGCATGTGACGAATGCAT 60
QY 436 atggcctaactccgggacaacaggtatggtctcagcagccgctgatatgctgataccca 495
|||||
Db 61 ATGGCTTAATCTCCGGACAGACAGGATGCTCAGCAGCGCTGTATGTGCTGATCCCA 120
QY 496 gcaaccccggaacacatgactgtgcaagcggttcacagagacatatctcctctc 555
|||||
Db 121 GCACCCCGGGAACCATAGCAAACTGTGCAAGCGGTTCCAGGACACATATATCTCTTC 180

QY 556 ccgctgcaacacatctccagtgactccgggaaacatagcgaatgagacatctccat 615
|||||
Db 181 CGCCTGCAOCCACATCTCCAGTACCTCCGGGAAACTAGCGAGATGAGACTCTCCAT 240
QY 616 cggctcctgacaagaatgtagaacaatgcaaatgtccglatltaagaacactgcgagca 675
|||||
Db 241 CGCCTCTCTGAGAAATGTATGAAACATGCAAGATTCCTGATTAAGAACCTCGAGCGAC 300
QY 676 caggaaacaatcagagagacacaggggtcagggtctgtgaaccccaagacgaacac 735
|||||
Db 301 CAGGAACACATCAGGAGACACACAGGCTTGTGAACCCAGAGACAGACAAAC 360
QY 736 agccttltgacatgactccacacactctcgcgtacacggttgatgagactgtcagact 795
|||||
Db 361 AGCCTTGTGACAGTATCCTCACACATCTCCGCTTACCGGAGATGTGAGTGTGCACAGT 420
QY 796 ctccatctgtgtgcaagttgtctcgtgctgtgctacacaacagcagttccagctact 855
|||||
Db 421 CTCCATCTGTGTGCCAGTGTCTCTGTGCTCTCACACAGGACAGTTCAGGCTACT 480
QY 856 acaacaatgaaatccggtatccctcggacagaacaagatccttctgcaactctccccc 915
|||||
Db 481 ACAATATGGAACATCCGGATTCCTGACAGCAACAGATCTTCTGACACTCTCCCC 540
QY 916 caggagccacttltgtlcaaggagacagccatgctagcaactccttgagacaacacagatcc 975
|||||
Db 541 CAGAGCCACTTGTGTGACAGGACAGGACATGCTGACATCTCTGAGCAACAGAGTCC 600
QY 976 ttcttgcaactcttcccccaagggtcaactltgtltaaggagacagccagctagcactc 1035
|||||
Db 601 TTTCTGTGCACTCTTCCCGCAGGGGTCACTTGTGTGACAGGACAGGCTTACGACATC 660
QY 1036 ctggagagacaagatccttctgcaactcttccccaggaatgcatctgtgtcaggagac 1095
|||||
Db 661 CTGGGACAGCAAGTCTCTTCTGCACTCTTCCCGAGAGTCACTTGTGTGACAGGAC 720
QY 1096 aggcacgccttagcaactctctggaacgaacaggtcttcttgcaacactctccagagag 1155
|||||
Db 721 AGGCCAGCGCTTACGACTCTGCGGACACAAAGGTCTCTTGTGACATCTTCCCGAGAG 780
QY 1156 ccaacttltgtcagatcaagatgatagtccttgaacacacacaggtctcgtgacaagagac 1215
|||||
Db 781 CCACCTTGTGTGATGATCAAGGATATCTGGAACATCCGAGATCTGTGACAGCAGGAGAC 840
QY 1216 agctagtgacagatgtgtgccccctcagatcccaacacctgcatgcgcgcacataga 1275
|||||
Db 841 AGCTATGTGACAGTGTGTGTGCCCCCTCAGATTTCAAAACCTGTGATCCGCCATTCATGA 900
QY 1276 acattagtggaatgggtatcctctctctacacgacatacagcccaaacctcgcatactgg 1335
|||||
Db 901 ACATATGTGGAATGGGTATCTCTTCTTACCCCATACAGCCCAAACTCGGATCACTCG 960
QY 1336 gatccctgtgtgacatacagaagaacggggggagacatctcgcgagcaaaacccgagaa 1393
|||||
Db 961 GATCTGTGTGTGACATACAGAAAGCGGGGAGACATCTCGAGCAAAAACCCGAGAA 1018

RESULT 6
AF310679 936 bp DNA linear INV 17-JUN-2001
LOCUS Encephalitozoon cuniculi III polar tube protein gene, partial cds.
DEFINITION AF310679
ACCESSION AF310679
VERSION AF310679.1 GI:14484927
KEYWORDS
SOURCE Encephalitozoon cuniculi.
ORGANISM Encephalitozoon cuniculi
Eukaryota; Microsporidia; Unikaryoniidae; Encephalitozoon.
REFERENCE 1 (bases 1 to 936)
AUTHORS Xiao,L., Li,L., Visvesvara,G.S., Moura,H., Didier,E.S. and Lal,A.A.
TITLE Genotyping Encephalitozoon cuniculi by Multilocus Analyses of Genes
with Repetitive Sequences
J. Clin. Microbiol. 39 (6), 2248-2253 (2001)
JOURNAL

```

MEDLINE      21270266
PUBMED      11376065
REFERENCE   2 (bases 1 to 936)
AUTHORS     Xiao,L., Li,L., Vivasvaram,G.S., Moura,H. and Lai,A.A.
TITLE       Direct Submission
JOURNAL     Submitted (04-OCT-2000) Division of Parasitic Diseases, National
            Center for Infectious Diseases, Centers for Disease Control and
            Prevention, 4770 Buford Highway, Atlanta, GA 30341, USA

FEATURES
    source
        1..936
            /organism="Encephalitozoon cuniculi"
            /isolate="CDC-V282"
            /specific_host="Homo sapiens"
            /db_xref="taxon:6035"
            /note="genotype: III"
            <1..>936
            /product="polar tube protein"
            <1..>936
            /codon_start=-3
            /product="polar tube protein"
            /protein_id="AAK62816.1"
            /db_xref="GI:14484928"
            /translation="ALMKLENYVSATALGSLNPQGGMAQPSYVLPSTPGTI
            ANCGSDPTSPSPAIPSPVPKPTSEMSPSPAPADVGTCKIAVLRKCPGTT
            SCTPSSGCEPTEPBOQPLSVISTTPAVPTVPSASPSVPPVAVAHKQAVPTN
            NGTSGIPGQDQIISGTLPPCATLCQGGQAMPSTGQDQVLSGTLPPGVTLQGGATPST
            PGQOQVLSGTLPPGATLCQDQMGPTSGVPGQGGQSGGQCAQIIPVPPSPMINS
            NGYPSSTAVSPNIGSLGSCVDIQTGTGTSCQKPE"

BASE COUNT   221 a      282 c      234 g      199 t

Query Match  46.2%; Score 844.8; DB 3; Length 936;
Best Local Similarity 92.1%; Pred. No. 5.5e-251;
Matches 934; Conservative 0; Mismatches 2; Indels 78; Gaps 1;

```

```

Db          541 GAGCCACTTGTGTACAGGACAGCCAT----- 568
Qy          979 ctggcactctcccccagggttaacttggtaaggagacagccagcctgaactctg 1038
Db          569 -----GCTAGCACTCTG 582
Qy          1039 ggcagcaacaggtctctctgtgacactctcccccaggagtaacttggtaaggagag 1098
Db          583 GGCAGCAACAGGTCCTTCTTGCACTCTCCCCAGGGGTCACTTGTGTGAGGACAGG 642
Qy          1099 ccaagcctagcactctctgggcagcaacaggtctctctgtgacactctccaggagca 1158
Db          643 CCAGCCCTAGCACCTCCGGGACGACACAGGCTCTTTCGGACCCCTTCCAGAGGCA 702
Qy          1159 cttgtgtcaagatcaaggtatgctgctggaacatccgagttccctggcagagagagat 1218
Db          703 CTTTGTGTACAGATCAAGGATGCTGCTGACATCCGGAGTTCCTGGACAGGACAGT 762
Qy          1219 ctatggacagtggtgtgccccctcagatcccaaacctgtcatgcgcacatcaaga 1278
Db          763 CTAGTGACAGTGTGGCCCCCTCAGATTCACAAACCTGTGATCGCGCATCATGAACA 822
Qy          1279 ttatggaaatgggtatccctctctctacgcatacaagccaaacctggatcaggat 1338
Db          823 TTAGTGAAATGGGTATCTTCTTACCCGATACAGCCCAACCTGGATCACTGGAT 882
Qy          1339 cctgtgtgaatacaagaacggggggagatactctgagcaaaacccgaga 1392
Db          883 CCTGTGTGACATACAGACAGGGGGGACATCTCTGGACAAACCCGAGA 936

RESULT      7
AX007051    1116 bp   DNA      linear   PAT 06-SEP-2000
LOCUS       AX007051
DEFINITION  Sequence 3 from Patent W00001724.
ACCESSION   AX007051
VERSION     AX007051.1 GI:9994977.
KEYWORDS
SOURCE      Encephalitozoon intestinalis.
ORGANISM   Encephalitozoon intestinalis.
REFERENCE   1 (bases 1 to 1116)
AUTHORS     Vivares,C., Danchin,A. and Delbac,F.
TITLE       Microsporidium polar tube proteins, nucleic acids coding for said
            proteins and their uses
JOURNAL     Patent: WO 0001724-A 3 13-JAN-2000.
            VIVARES CHRISTIAN (FR); DANCHIN ANTOINE (FR); DELBAC FREDERIC (FR);
            CENTRE NAT RECH SCIENT (FR)
            Location/Qualifiers
                source
                    1..1116
                        /organism="Encephalitozoon intestinalis"
                        /db_xref="taxon:58839"
BASE COUNT   300 a      332 c      230 g      254 t
ORIGIN
Query Match  9.6%; Score 176; DB 6; Length 1116;
Best Local Similarity 53.5%; Pred. No. 6.2e-43;
Matches 635; Conservative 0; Mismatches 460; Indels 93; Gaps 9;

```


RESULT	8
AF044915	
LOCUS	1696 bp DNA
DEFINITION	linear INV 26-AUG-1998
	Encephalitozoon hellem polar tube protein PTP5 precursor (PTPH5)

Oy 1389 gagaagtcgcgcacgcatatgccataggagcctgtgaacaaccaacaagaattat 1448
DB 1345 GAAAGTCTACTATGTCACATATGCCATGTGGAGCCTGTGCGGCACCACCAACCCCAACGTTTTC 1404

Oy	1449	ataggcaacagc	gagat	actctgttg	gacacga	aagtaca	gaat	aactctc	atgc	1508	
Db	1405	ataggtaaac	gagga	atattct	ctgtcg	ggcgagg	aatg	atcact	ctccatgc	1464	
Oy	1509	aacactgtct	gccaa	tgtctgt	cagcta	ataata	aagagt	taatc	tctcttc	ctcg	1568
Db	1465	aaactctct	ctgtcaa	---	tgttag	cgctaa	aaataa	atgatt	aaactct	ctctaa	1521
Oy	1569	gtctt	1573								
Db	1522	gtttt	1526								

[illegible]

REFERENCE	1 (bases 1 to 1340)
AUTHORS	Xiao, L., Li, L., Moura, H., Sulaiman, I., Lal, A.A., Gatti, S., Scaglia, M., Didier, E.S. and Visvesvara, G.S.
TITLE	Genotyping <i>Encephalitozoon hellem</i> isolates by the Polar Tube Protein Gene
JOURNAL	J. Clin. Microbiol. 39 (6), 2191-2196 (2001)
PUBMED	11376056
REFERENCE	2 (bases 1 to 1340)
AUTHORS	Xiao, L., Li, L., Moura, H., Lal, A.A., Gatti, S., Scaglia, M. and Visvesvara, G.S.
TITLE	Direct Submission
JOURNAL	Submitted (19-JAN-2001) Division of Parasitic Diseases, Centers for Disease Control and Prevention, 4770 Buford Highway, Atlanta, GA 30351, USA

FEATURES	
source	Location/Qualifiers
	1..1340
	/organism="Encephalitozoon hellem"
	/isolate="CDC:V261"
	/specific_host="Homo sapiens"
	/db_xref="taxon:27973"
	/note="genotype: 2B"
mRNA	<1..>1340
CDS	/product="polar tube protein"
	<1..>1340

```

/codon_start=3
/product="polar tube protein"
/protein_id="NAK63046.1"
/db_xref="GI:14485633"
/translation="MKLGNYSAVPLCSNTYDPSQQQPSYLVLPSTPEALINCAVSPC
NAVYSPSTTSSPTGNTNDNESPTEEDVOTCKISVYKHDFGASSTPCEPEPTIILP
AQPPTMATVTPALLASVOTPESVAVIPYTOAVIQPAPITVPSPIIPANQAIKGYAPV
GTFPAPGQGGIILSGSYLAPGASSCOLVPGMTPLGMPGTGPIGLYPCITGGGSGNS
TIPPIVPCPGONGDGSNORTIPGIVPCPGONGSGSSGNTTGGIYPCPGONGS
GSGSNORTIPGIVPCPGONGSGSGSNNTTGGIYPCPGONGSGSNNTTGGIYPC
PGONGSGSGNNTTGGIYPCPGONGSGSSSSTIPGVISPCPGONGSGNNGTGGD
GQCVSYDQTPPLAMPPLTSGISGNGPSTSTYTGSLQGLGCPIDVQKPTSNCS"
372 a 354 c 314 g 300 t

```

Query Match	5.6%	Score 102.6	DB 3	Length 1340
Best Local Similarity	66.9%	Pred. 10.4	4.4e-20	
Matches	204	Conservative	0	Mismatches 74; Indels 27; Gaps 3.
OY	385	tgatgaagttgagaagtctatcttaagcaaccgcgactgtgcagcaatgcatatgacctaa	444	
Db	1	TAATTAACCTGGGGCAATGTTTATTTGTGCGCTTCCCTTGGCGACATATCTTTGATTCCT	60	

OY 445 ctcgcggacaacaggtagtgcctcagaagccgctgtatgtctataccccagaaccgccg 504
| | | | | | | | | | | | | | | | | |
Db 61 C-----GCACAGCAGCCATCATATGTGCTGGTTCCTCGACTCTTG 102

OY 505 gaaccatagcaactgycgaagcggttacagagacatatctctctccgcgtcac 564
| | | | | | | | | | | | | | | | | |
Db 103 AGGCCATTAAACAAMCTGTCATPACACCCECAAGAATGCAATATGTCCCTTTCTCTCACTA 162

OY 565 ccacattccagttgacctccggaaaaactaagaaatagagaaactctccatccgctctg 624
| | | | | | | | | | | | | | | | | |
Db 163 CATCTATCTTCA---ACCCAGGAGAACMAACGAATACGAACATCTC-----CTACTA 213

OY 625 cagaagatgttagaaacatgccaagtctcgatattgaagcaactgcgcgcaccagaabaac 684
| | | | | | | | | | | | | | | | | |
Db 214 CAGAGAGATGTAGGAACATGCACGAATTTCCGTTGTAAACATTTGCACACACCCGGGTCAT 273

OY 685 catca 689
| | | | |
Db 274 CATCA 278

RESULT	10			
AF338361				
LOCUS				
DEFINITION	AF338361	1184 bp	DNA	linear INV 18-JUN-2001
ACCESSION	Encephalitozoon hellem	genotype 1A	polar tube protein gene, partial cds.	
VERSION	AF338361			
KEYWORDS	AF338361.1	GI:14485648		
SOURCE				
ORGANISM	Encephalitozoon hellem.			
REFERENCE	Encephalitozoon hellem			
AUTHORS	Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.			
	1 (bases 1 to 1184)			
	Xiao, L., Li, L., Moura, H., Sultman, I., Lal, A. A., Gatti, S.,			
	Scaglia, M., Didier, E. S. and Vivasvora, G. S.			
	Genotyping Encephalitozoon hellem isolates by Analysis of the Polar			
	Tube Protein Gene			
	J. Clin. Microbiol. 39 (6), 2191-2196 (2001)			
TITLE				
JOURNAL				
PUBMED	11376056			
REFERENCE	2 (bases 1 to 1184)			
AUTHORS	Xiao, L., Li, L., Moura, H., Lal, A. A., Gatti, S., Scaglia, M. and			
	Vivasvora, G. S.			
	Direct Submission			
TITLE	Submitted (19-JAN-2001)			
JOURNAL	Disease Control and Prevention, 4770 Buford Highway, Atlanta, GA			
	30341, USA			

FEATURES	source	Location/Qualifiers
		1..1184
		/organism="Encephalitozoon hellem"
		/isolate="CDC:V274"
		/specific_host="Homo sapiens"
		/db_xref="taxon:27973"
		/note="genotype: 1A"
		<1..>1184
mRNA		/product="polar tube protein"
		<1..>1184
CDS		/codon_start=1
		/product="polar tube protein"
		/protein_id="AAK63044.1"
		/db_xref="GI:14485649"
		/translation="STWVKLGNYSAVPLCSNTYDPSQQQPSYVILIPSPPEAITNCA VSPNNAVYSSPTTSSSTPTGNNNDNETSPTEDEGTCISKVKKCDPAGASSTPCBE QTIAPQPTATVATPAIIASVQPTSPSVVLPEDKVIQAPATMIVPSSIIIPGYNPE PAIGQOGLISGSGVLAPGASCOLVPGNTPGQMLPGMTGVSCLPTGGGDSNQT IPIVGPQGGGGSGSNQTIPIGVSPQGGGGSGSNQTIPIGVPCQBPQGGGSGNSQ TIRIVISPCQBPQGGGSSNQTIRIGIYVPCQBPQGGGDSNQTIRIVISPCQBPQGGGN GTTQPGQCVSPQTPNPIAMPPIISGNGIPSTYITSLGQLGCPIDVQKPTSSC ES"
BASE COUNT	314 a	321 c
ORIGIN	274 g	275 t

Query Match 5.58; Score 101.4; DB 3; Length 1184;

BASE COUNT	325 a	337 c	290 g	280 t
ORIGIN	GQLPCPCIDVOKPFISSCS"			
Query Match	5.4%;	Score 99;	DB 3;	Length 1332;
Best Local Similarity	66.3%;	Pred. No. 5.6e-19;		
Matches 201;	Conservative	0;	Mismatches 75;	Indels 27; Gaps 3;
Oy	367	atgaagcttgaggaaatgctattcttcagcaacccgacgtgctcagcaatgcatatgctctaact	446	
Dd	2	ATGAAGCTGGGGAAAGTTTATCTCAATTCCGCTTTCAGTAATACCTTAATGATCCCTC-	60	
Oy	447	ccgggacaacacagggatgctcgaagacgcgtcgtatgctcgtatcccaagcaccccgga	506	
Dd	61	-----GCAACAGCAGCCATCATATGTCGATGCCAGATACCTCTAG	103	
Oy	507	accataagcaaacgtgcaagcggttcacaggaacatatcttccttcgcgctgaccc	566	
Dd	104	GCTATMACAAACTGTGCATACAGCCCAAGAAATGCAATGTGCCCTTCTTCCCTACTAC	163	
Oy	567	acatctcgaagctacccggggaacactggcgaaatgtagatcttccttcgctctctga	626	
Dd	164	TCGCTTTCA--ACCCGAGGAGCAAAACAGTAACGAAACATCTC-----CTACTACA	214	
Oy	627	gaagatctgaggaacatgcaagatctgcgtatggaacactgcgaagcaccccggaaca	686	
Dd	215	GAGGATGTGAGCAACTGCATGCAAGATTCCCTTGTAAAGCATTCGCAACACACAGTGCATCA	274	
Oy	687	tca	689	
Dd	275	TCA	277	
RESULT 13				
AF429315/c				
LOCUS	AF429315	125020 bp	DNA	linear PRI 18-JAN-2002
DEFINITION	Homo sapiens junctophilin 3 (JPH3) gene, partial cds.			
ACCESSION	AF429315			
VERSION	AF429315.1	GI:17646244		
KEYWORDS	human.			
SOURCE	Homo sapiens			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 125020)			
AUTHORS	Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,			
	Ingersoll,Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,			
	Potter,N.T., Ross,C.A. and Margolis,R.L.			
TITLE	A repeat expansion in the gene encoding junctophilin-3 is			
	associated with Huntington disease-like 2			
JOURNAL	Nat. Genet. 29 (4), 377-378 (2001)			
MEDLINE	21583737			
PUBMED	11694876			
REFERENCE	2 (bases 1 to 125020)			
AUTHORS	Holmes,S.E., Ingersoll,Ashworth,R.G., Ross,C.A. and Margolis,R.L.			
TITLE	Direct Submision			
JOURNAL	Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical			
	Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA			
FEATURES	Location/Qualifiers			
source	1..125020			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/chromosome="16"			
	/map="16q24.3; between D16S520 and WT-12410"			
	/note="Isolated from a patient with Huntington's			

	repeat_region	Disease-Like 2 (HDL2)" complement(35581.._35746) /rpt_type= tandem /rpt_unit=ctg complement(<35507..>36687)
	mRNA	/gene="JPH3" /product="junctionhilin 3" complement(<35507..>36687) /gene="JPH3" /note="JPH3"
	gene	complement(<35507.._36687) /gene="JPH3" /note="component of the junctional complex between plasma membrane and endoplasmic reticulum"
	CDS	/codon_start=1 /product="junctionhilin 3" /protein_id="AAI0941.1" /db_xref="GI:17646245" /translation="MSSGGRENFDGSGYCGGWEDGAHGVCGRPQGGEGYSNWS HGFEVLGYTPWPSGNTYOGTGAOCGRIGILESKGVVYGEMTHGPKRGYVEKAG NGARIEGWSMGLODGGTYESDG"
	BASE COUNT	29056 a 32731 c 30696 g 28283 t 4254 others
	ORIGIN	
	Query Match.	2.7%; Score 49.4; DB 9; Length 125020;
	Best Local Similarity	12.2%; Pred. No. 0.0043;
	Matches 121;	Conservative 424; Mismatches 442; Indels 8; Gaps 3;
Oy	525 agcggatcacaggaacacatatcccttcctccgcgtgcaaccacacatcccaagtacccg	584
:::	::: :: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	51634 WKSMTWTGMRRMGTERRAMMMWWKKYYRBMKCATRSMWMAASKSYARKCCSMM	51575
Oy	585 gggaactagagagaatagagacatcccatcgctctctgccaagaatgtaggaaatgc	644
:::	::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	51574 AYKRYKSRASTRCKGKGMAKSCMRMYMYKWGGGAGKCCSAAGMTGRYSRASCCMKMS	51515
Oy	645 aagattgcgatlttgaagcatcgagccagcaagaacaacatacagaggagaccagg	704
:::	::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	51514 KKSASAKAYAVASTDAAMAAMSGCTCYMKCTPRMAATACMARRMWAAMWSCMRKCTWSK	51455
Oy	705 tcagggcttgtlaaacccccagagagcaaacagaccttgta-qtgattccaccactcc	763
:::	::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	51454 MSKMGRMRHYCWSMARMMYYLAGSYNSKCRSRMGYKGRSRSHSGMSWTRGYSARCYM	51395
Oy	764 tgcgtaccggtgtagctgtgagagctgcacagctccatctgctgtgtgccagtgtgctcgt	823
:::	::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	51394 SKSAGAGACKCKAGSYANGSAGCMRWGSYKWMYMSCRSMAKSKSMTCYTWSCCTCSKGC	51335
Oy	824 cgltgctaccaccaagcgagctcagagctcagctactacaacaatgnaaatccggtatccgtg	883
:::	::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	51334 YMYMSWKYGAOMCYMRMCCTCMARMMRRMAAARRRMAAGAARAWMSMMMSMMMR	51275
Oy	884 acagaacaagatccttctgagctcttcccacaggaagcactttgtlcaaggagcagc	943
:::	::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	51274 BHAALMALRMMAALACMMWYSYRKTCWMSMAAAMAMMWMCIRAAWYKKTETMBRRS	51215
Oy	944 catgctcagcactcctgagcagcaacagatcc---ttctgtgcaactcttccccagg	999
:::	::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	51214 KYSSSYAAWAMYTMWCYSSMYAARMGTYSGKRKRMASTSCMCCRRCCKMKKKGR	51155
Oy	1000 tcaactttgtlcaaggagcaagccactagacatccctcgggagcaaaaagtgctctctg	1059
:::	::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	51155 WCSTTTTGHSKGSKSWHWISSMACMTWYGRSKRTTSMATSGCMRRGAMSGMKMKYVSC	51095
Oy	1060 gaactcttcccacagagatcaactttgtlcaaggagcagccagcactagcactctcgtgc	1119
:::	::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	51094 RKKGMRWGWSYKCCSGYMSMGMTGSGTSKYCYTGGYCSCAKXKCKSKSTSCKCT	51035
Oy	1120 agcaacaggtccttctgtgcaacccttccccaaggagcactttgtgtcaagataagta	1179
:::	::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	51034 KSKYCSGMRGSSYCTCGWGRCKMKCMSKSCSRSGCKSYKMAAGSMRCRGRGMSGMTWS	50975

Oy	1180	tgccggagacatccgagatctctgtgcaacgagcagctctgtgtggaacgttgtgccc	1239
Db	50974	SMKSTSMASYKSASANGRGKCTKCTSCMTYYTMMHKSSTHDSBGYKSYKSYWTGCMKMD	50915
Oy	1240	ctcagatcccaaacctcttcattgcccgcattccatgacaattagttggaatgggtacctt	1299
Db	50914	SMCWGVDSHMMRDMMYDKMKRMKVSSTGTYGHHBSHSMCMKMSRSDSDSGYGMHMKSYWK	50855
Oy	1300	ctctaccgcatacaagcccaaacctccgcgatactgtggtacctgttctgtgacataacgaaga	1359
Db	50854	SSASHHMTYSGMKSSVYBDDCSMTTSBSKSYRMRRRSGSTBYRKCKAVMRYHRST	50795
Oy	1360	cggggggggacatccctgcgcagacaanaac-----cgaagaatccgcacgcagatgacatg	1416
Db	50794	RSVMTGGRKSCVVDVSYSRRBSVHTSMSSWMRCRCRGARBMKSCAGKMSCVMTGMCVHSR	50735
Oy	1417	aggcctgtgtcaacacccaacacacacggtattattatgagcaacagcgagtatctgttgac	1476
Db	50734	CGSMKCMSCDDCCRRABDAHVCPCBMYKBDACBCKYRDSVSCSAHVMVWVBHBSKS	50675
Oy	1477	caggaatgataactgcaattactcttcattctgcgaac	1511
Db	50674	SRBGWRHMHACHKSMCKKSYVKNVHVKMKCYWC	50640
RESULT	14		
LOCUS	AF429315		
DEFINITION	Homo sapiens junctophilin 3 (JPH3) gene, partial cds.		
ACCESSION	AF429315		
VERSION	AF429315.1	GI:17646244	
KEYWORDS	human.		
SOURCE			
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 125020)		
TITLE	Holmes,S.E., O'hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S., Ingersoll-Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A., Potter,N.T., Ross,C.A. and Margolis,R.L.		
JOURNAL	A repeat expansion in the gene encoding junctophilin-3 is associated with Huntington disease-like 2		
PUBMED	Nat. Genet. 29 (4), 377-378 (2001)		
REFERENCE	2 (bases 1 to 125020)		
AUTHORS	Holmes,S.E., Ingersoll-Ashworth,R.G., Ross,C.A. and Margolis,R.L.		
TITLE	Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical Institutions, 600 N. Wolfe St., Baltimore, MD 21207, USA		
JOURNAL			
FEATURES			
Source	Location/Qualifiers		
	1..125020		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="16"		
	/map="16q24.3: between D16S520 and WT-12410"		
	/note="Isolated from a patient with Huntington's Disease-Like 2 (HDL2)"		
	complement(3581..35746)		
	/rpt_type=tandem		
	/rpt_unit=ctg		
	complement(<36507..>36887)		
	/gene="JPH3"		
	/product="Junctophilin 3"		
	complement(<35507..>36887)		
	/gene="JPH3"		
	/note="Jp3"		
	complement(<36507.. 36887)		
	/gene="JPH3"		
	/note="component of the junctional complex between plasma membrane and endoplasmic reticulum"		
	/codon_start=1		
	/product="Junctophilin 3"		

Query Match	Best Local Similarity	2.6%	Score 47.4	DB 9	Length 125020
Matches 76	Conservative 286	Mismatches 298	Indels 4	Gaps 2	
Query 464	ggcctagcagccgctgtagtgcgtatgcctcccccagcagcccgagggaaccatagcaaatgtgc	523			
Db 16900	SVSYVKHMSMASBSCHMBKCMWVSCMSMMYKSSMMGSSWGMCMGRRRSKMKWYS	16959			
QY 524	aagcggttcacagaaacatatctctctctccctgcgtgcacacatctccagtgtaccc	583			
Db 16960	RCMSRSKSMRYTGGSKMMSSMCTSSCYASMCMCMSCCMRSCCCRCRYCCCMRYC	17019			
QY 584	ggggagaaatagcaggaatgaacatctccatcgcgtcctctgcagaagaatgtggaacatg	643			
Db 17020	CACKCYMSSYTWMSASYSMSRYSKMYKMSKMSKMSRMSKRSKSKSGGSGSKGKGGS	17079			
QY 644	caagattgcgcatgtaagcaactgcagccagcagcaagaacaataggaagagacacag	703			
Db 17080	YGRKTKRKSRCGM - KGAKMYMYRRSMRKMKYSSKGYMCMYCMWGRGCGSCMTSRS	17138			
QY 704	gtcagggcctgtgaaacccacagacagcaacagccttgcattgcatgtatctccacatcc	763			
Db 17139	AMCSCYCAKCKSMCYSCYGGSMKGYAACRYRCSMSKSKYCMRSTYSTSCGCCCTTYY	17198			
QY 764	tgcgctacgcgtgactgtgagtcgtgcacagctcccatctgtgtgccaagtgtctcgt	823			
Db 17199	TCCCCCNANTGGGGAAGCTTTNCMKKITYRYRKRNKGCAMCYNNYNNMSRSCRAGMS	17258			
QY 824	cggtgcacacacagcaagtcacagctactacaaatggaacatccggtatctctg	883			
Db 17259	CTKYKSSMTMSASACWCMSSMYKCSMRSAISGMSWSYMKMKMSWMSRSCMYCMK	17318			
QY 884	aaagaaacagatccctctgcacctctcccccagagagcacttgtgtcagggacag--	941			
Db 17319	MCYCMRSMRSGMSTYMAKSSSRGCTCRCYWCMSKSCYASYVMMSKMKMKK	17378			
QY 942	gcacatgcctagcactctgcagacaaacagatccctctgcacactctcccccaggg	1000			
Db 17379	WRCMGSMWGSASRSRYCYKCYKSMKMSKSCYACAGOMMKGGYMYMYRMSWKRRMW	17438			
QY 1001	caacttgcgtcaggaagcagcagcagcactgcactctgtgcagcaacagtcctctctg	1060			
Db 17439	GKSAMYRMRMRMKGGANMMCKYCSRMSRCMMKSYCASCRSCAMMSGYMKCASYC	17498			
QY 1061	caactctcccccagagatcaacttgcataaggaagcagcagcctagcactccctggca	1120			
Db 17499	MSSYCTCWCISYMYNCTCCKCKMSSYTKSKSSMSSTKCKRKRKSYSCMCTSRGAM	17558			
QY 1121	gcaa 1124				
Db 17559	SCMR 17562				
RESULT 15					
CNS05FC3					
LOCUS	CNS05FC3	203108 bp	DNA	linear	PRI 13-NOV-2001
DEFINITION	Human chromosome 14 DNA sequence BAC R-79666 of library RPCT-11				
ACCESSION	AL155032				
VERSION	AL155032.6	GI:16944347			
KEYWORDS	HTG.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: September 1, 2002, 00:28:16 : Search time 303.84 Seconds
(without alignments)
10340.816 Million cell updates/sec

Title: US-09-755-456-1

Sequence: 1 gaattcagatgcctcatacc.....tccctagaagtgtggtac 1830

scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Listing first 45 summaries

Database : N.Genseq_032802.*

1:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1980.DAT.*
2:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1981.DAT.*
3:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1982.DAT.*
4:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1983.DAT.*
5:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1984.DAT.*
6:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1985.DAT.*
7:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1986.DAT.*
8:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1987.DAT.*
9:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1988.DAT.*
10:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1989.DAT.*
11:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1990.DAT.*
12:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1991.DAT.*
13:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1992.DAT.*
14:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1993.DAT.*
15:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1994.DAT.*
16:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1995.DAT.*
17:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1996.DAT.*
18:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1997.DAT.*
19:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1998.DAT.*
20:	/SIDS/gcgdata/genseq/genseqn-emb1/NA1999.DAT.*
21:	/SIDS/gcgdata/genseq/genseqn-emb1/NA2000.DAT.*
22:	/SIDS/gcgdata/genseq/genseqn-emb1/NA2001A.DAT.*
23:	/SIDS/gcgdata/genseq/genseqn-emb1/NA2001B.DAT.*
24:	/SIDS/gcgdata/genseq/genseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	DB	ID	Description
1	1830	100.0	1830	21	AAZ91192	Encephalitozoon cu
2	176	9.6	1116	21	AAZ91194	Encephalitozoon cu
3	45.4	2.5	740	22	AAAF45087	Human secreted pro
4	38.2	2.1	611	22	ABA47116	Human breast cell
5	38.2	2.1	611	22	ABA64998	Human foetal liver
6	38.2	2.1	611	22	ABA32106	Probe #10572 for g
7	38.2	2.1	611	22	AAK13424	Human brain expres
8	38.2	2.1	611	22	AAK31961	Human bone marrow
9	38.2	2.1	611	22	AAI19970	Probe #9903 for ge

C	45	34.2	1.9	1166	20	AAI19440
C	44	34.2	1.9	444	22	AAK60027
C	43	34.4	1.9	10369	24	ABL132332
C	42	34.4	1.9	10083	23	AA546303
C	41	34.4	1.9	10083	23	ABL12194
C	40	34.4	1.9	2113	22	AAFC88071
C	39	34.8	1.9	4182	22	AAFS8998
C	38	34.8	1.9	4182	21	AAI222758
C	37	34.8	1.9	4182	20	AAAX03945
C	36	34.8	1.9	4182	20	AAAX25388
C	35	34.8	1.9	4182	20	AAAX25388
C	34	34.8	1.9	3620	18	AAV06054
C	33	34.8	1.9	3448	21	AAH08581
C	32	34.8	1.9	3448	21	AAH08581
C	31	34.8	1.9	3448	20	AAAX03946
C	30	34.8	1.9	3448	20	AAAX25389
C	29	34.8	1.9	3448	20	AAAX25389
C	28	35	1.9	4590	23	AAH24065
C	27	35.2	1.9	9947	23	ABL09444
C	26	35.4	1.9	9577	22	AAK80931
C	25	35.6	1.9	2944538	24	ABAO04214
C	24	35.6	1.9	12235	23	AAH33069
C	23	35.6	1.9	8269	22	AAAS3346
C	22	35.6	1.9	8269	22	AAAS3346
C	21	35.6	1.9	8269	22	AAAS1976
C	20	35.6	1.9	8269	22	AAAS1472
C	19	35.6	1.9	8269	22	AAI64010
C	18	35.6	1.9	8269	22	AAK90941
C	17	35.6	1.9	8269	22	ABA06793
C	16	35.6	1.9	8269	22	AAI99642
C	15	36	2.0	1090	22	AAI996405
C	14	37.2	2.0	2880	10	AAAN92068
C	13	37.2	2.0	2880	10	AAAN92068
C	12	38.2	2.1	2465	11	AAO02068
C	11	38.2	2.1	611	22	AAI45168
C	10	38.2	2.1	611	22	AAI45168
C	9	38.2	2.1	611	22	AAI45168
C	8	38.2	2.1	611	22	AAI45168
C	7	38.2	2.1	611	22	AAI45168
C	6	38.2	2.1	611	22	AAI45168
C	5	38.2	2.1	611	22	AAI45168
C	4	38.2	2.1	611	22	AAI45168
C	3	38.2	2.1	611	22	AAI45168
C	2	38.2	2.1	611	22	AAI45168
C	1	38.2	2.1	611	22	AAI45168
C	0	38.2	2.1	611	22	AAI45168

ALIGNMENTS

RESULT	1
AAZ91192	
ID	AAZ91192 standard; DNA; 1830 BP.
XX	
AC	AAZ91192;
XX	
DT	21-JUN-2000 (first entry)
DE	Encephalitozoon cuniculi polar tube protein PTP55 gene.
XX	
KW	Vaccine; gene therapy; polar tube protein; microsporidium; antibody;
KW	antiparasitic; antiprotozoan; immunostimulatory; immunosassay;
KW	opportunistic pathogen; immunocompromised patient; AIDS; ds.
XX	
OS	Encephalitozoon cuniculi.
XX	
PN	MO200001724-A1.
XX	
PD	13-JAN-2000.
XX	
PF	06-JUL-1999; 99MO-FR01630.
XX	
PR	07-JUL-1998; 98FR-0008692.
XX	
PA	(CNRS) CNRS CENT NAT RECH SCT.
XX	
PI	Delbac F, Vivares C, Danchin A;
XX	
DR	WPI; 2000-237186/20.
XX	
PT	P-PSDB; AAY83760.
XX	
XX	New microsporidial tube protein and related nucleic acid, useful e.g

PT in vaccines against Encephalitozoon infection and for raising
XX diagnostic antibodies -
PS Claim 17; Page 39-41; 55pp; French.

XX This sequence represents the complete gene encoding the polar tube
CC protein pPp55 from the microsporidium Encephalitozoon cuniculi.
CC The protein has antiparasitic, antiprotozoal and immunostimulatory
CC activities. The proteins are used to raise specific antibodies, useful
CC as immunoassay reagents for detecting serum antibodies against
CC Encephalitozoon (opportunistic pathogens in immunocompromised,
CC especially acquired immune deficiency syndrome, patients), in
CC vaccines to prevent such infections.

XX Sequence 1830 BP; 460 A; 486 C; 434 G; 450 T; 0 other;

Query Match 100.0%; Score 1830; DB 21; Length 1830;
Best Local Similarity 100.0%; Pred.No.0;
Matches 1830; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gaattcagatgctcattacattggaatlaaaattatgattcatgttatatctg 60
Db 1 gaattcagatgctcattacattggaatlaaaattatgattcatgttatatctg 60
QY 61 ggcggagacggcgctcgtatcttctcaagggtgtcgctcaaccagtgacaggagttcc 120
Db 61 ggcggagacggcgctcgtatcttctcaagggtgtcgctcaaccagtgacaggagttcc 120
QY 121 ggggggtctctggtggaagaaagtcattgtgtgtctctacacatgctacgtccct 180
Db 121 ggggggtctctggtggaagaaagtcattgtgtgtctctacacatgctacgtccct 180
QY 181 ttcgctgtttccacaaagatcccaattatctctccaggaacttcaaccctcagaatgaa 240
Db 181 ttcgctgtttccacaaagatcccaattatctctccaggaacttcaaccctcagaatgaa 240
QY 241 cagaagatgaaactctctgtgcaaatcgtatagatcgtattgagacatttgaaccagag 300
Db 241 cagaagatgaaactctctgtgcaaatcgtatagatcgtattgagacatttgaaccagag 300
QY 301 ttggaataaagatataaatactctccgaaacgacaggtttaagatgaaggtatttcta 360
Db 301 ttggaataaagatataaatactctccgaaacgacaggtttaagatgaaggtatttcta 360
QY 361 agatcctctcgtcctatattgcccctgataagttggaagatgtctatccagcaaccgac 420
Db 361 agatcctctcgtcctatattgcccctgataagttggaagatgtctatccagcaaccgac 420
QY 421 tctgacagcaatgcatatgacctaaactccgggacaaagggtatgtctcagcagcgt 480
Db 421 tctgacagcaatgcatatgacctaaactccgggacaaagggtatgtctcagcagcgt 480
QY 481 atgtgtgattcccccagaccccgggaaacatagcaaacgtgtgcaacggttccagagaa 540
Db 481 atgtgtgattcccccagaccccgggaaacatagcaaacgtgtgcaacggttccagagaa 540
QY 541 catattctcctctcccgctgacccacacatctccagtgactccggggaaacatagagaa 600
Db 541 catattctcctctcccgctgacccacacatctccagtgactccggggaaacatagagaa 600
QY 601 atgagagcatctcattcgtctctctgcaagaatgtagaacaatgcaagattgcccattga 660
Db 601 atgagagcatctcattcgtctctctgcaagaatgtagaacaatgcaagattgcccattga 660
QY 661 agcacgtcgacagcagcaagaaacacatcagggagacacacaggggtcaggcccttgagaa 720
Db 661 agcacgtcgacagcagcaagaaacacatcagggagacacacaggggtcaggcccttgagaa 720
QY 721 ccccgagagcagcaagcctttgtcagtgatctccacacactcctcgtcagacaggtgactg 780
Db 721 ccccgagagcagcaagcctttgtcagtgatctccacacactcctcgtcagacaggtgactg 780

QY 781 tggagctcagcagctccatctgtgtgcccagttgtctcgtctgtctcaccacagg 840
Db 781 tggagctcagcagctccatctgtgtgcccagttgtctcgtctgtctcaccacagg 840
QY 841 cagttccagctactacaataatggaacatccgattctcttgaaacagcaacagatctct 900
Db 841 cagttccagctactacaataatggaacatccgattctcttgaaacagcaacagatctct 900
QY 901 ctgacactctcccccagagccactttgttaaggagacagcagctagtaactcttg 960
Db 901 ctgacactctcccccagagccactttgttaaggagacagcagctagtaactcttg 960
QY 961 gacagcaacagatccttctgtgacactctcccccaggggtacattgtgtcaggagag 1020
Db 961 gacagcaacagatccttctgtgacactctcccccaggggtacattgtgtcaggagag 1020
QY 1021 ccacgctcagcactcctctgagcaacagagtccttcttgacactctcccccagagta 1080
Db 1021 ccacgctcagcactcctctgagcaacagagtccttcttgacactctcccccagagta 1080
QY 1081 cttgtgtcagagcagcagccacgctcagcactcctctgagcaacagagtccttcttgca 1140
Db 1081 cttgtgtcagagcagcagccacgctcagcactcctctgagcaacagagtccttcttgca 1140
QY 1141 ccctctcccaaggagccactttgtcagagatcaaggatagtccttggaacatccggatcc 1200
Db 1141 ccctctcccaaggagccactttgtcagagatcaaggatagtccttggaacatccggatcc 1200
QY 1201 ctgagacagcagcagcagctcagtgagcagtggtgtgcccctcagatctcaaacctgtca 1260
Db 1201 ctgagacagcagcagcagctcagtgagcagtggtgtgcccctcagatctcaaacctgtca 1260
QY 1261 tgcgcacatccatgaaacatagtgagaaatggtatcctctctacacgacatccagccaa 1320
Db 1261 tgcgcacatccatgaaacatagtgagaaatggtatcctctctacacgacatccagccaa 1320
QY 1321 acctggatcaactggatcctctgtgtgacatacagaacagggggagacatcctgagac 1380
Db 1321 acctggatcaactggatcctctgtgtgacatacagaacagggggagacatcctgagac 1380
QY 1381 aaaaaccgagaaagttccgcagcagatagccaatgagacccgtgtgaaacacacaccaa 1440
Db 1381 aaaaaccgagaaagttccgcagcagatagccaatgagacccgtgtgaaacacacaccaa 1440
QY 1441 cggattatagcacaacagaaatcctgtgtgacacaggaatgtaacatgcaatgaact 1500
Db 1441 cggattatagcacaacagaaatcctgtgtgacacaggaatgtaacatgcaatgaact 1500
QY 1501 ctccatgcaaacactgctgtccaaatgctgtaggtctaaataaagagtttaactctct 1560
Db 1501 ctccatgcaaacactgctgtccaaatgctgtaggtctaaataaagagtttaactctct 1560
QY 1561 tttctcgtctcttttgaaagctttggaatgggagtgaggggtctatagggctgaagtgaat 1620
Db 1561 tttctcgtctcttttgaaagctttggaatgggagtgaggggtctatagggctgaagtgaat 1620
QY 1621 gccaaacactctctcgtcccaagaacacatctcgatgttctcctgtgagcagagtttgg 1680
Db 1621 gccaaacactctctcgtcccaagaacacatctcgatgttctcctgtgagcagagtttgg 1680
QY 1681 taacaggaatcccccagagattagcagccttgagagtaacatgatacagtaataaac 1740
Db 1681 taacaggaatcccccagagattagcagccttgagagtaacatgatacagtaataaac 1740
QY 1741 tttcctaattattattttcttctgtttatattcccgagccaactctggagaagaatgcttc 1800
Db 1741 tttcctaattattattttcttctgtttatattcccgagccaactctggagaagaatgcttc 1800
QY 1801 gaattcaagctcccttagaagtggtgagtc 1830
Db 1801 gaattcaagctcccttagaagtggtgagtc 1830

```

RESULT 2
AA291194
ID AA291194 standard; DNA: 1116 BP.
XX
AC AA291194:
XX
DT 21-JUN-2000 (first entry)
XX
DE Encephalitozoon intestinalis polar tube protein PTP55 gene.
XX
KM Vaccine; gene therapy; polar tube protein; microsporidium; antibody;
KW antiparasitic; antiprotozoan; immunostimulatory; immunosassay;
XX opportunistic pathogen; immunocompromised patient; AIDS; ds.
OS Encephalitozoon intestinalis.
XX
PN MO200001724-A1.
XX
PD 13-JAN-2000.
XX
PF 06-JUL-1999; 99MO-FR01630.
XX
PR 07-JUL-1998; 98PR-0008692.
XX
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI Delbac F, Viveres C, Danchin A;
XX
DR MPI: 2000-237186/20.
XX
DR P-PSDB; AA183762.
XX
PT New microsporidial tube protein and related nucleic acid, useful e.g.
PT in vaccines against Encephalitozoon infection and for raising
PT diagnostic antibodies
XX
PS Claim 18: Page 43-44; 55pp; French.
XX
CC This sequence represents the complete gene encoding the polar tube
CC protein PTP55 from the microsporidium Encephalitozoon intestinalis.
CC The protein has antiparasitic, antiprotozoal and immunostimulatory
CC activities. The proteins are used to raise specific antibodies, useful
CC as immunoassay reagents for detecting serum antibodies against
CC Encephalitozoon (opportunistic pathogens in immunocompromised,
CC especially acquired immune deficiency syndrome, patients), in
CC vaccines to prevent such infections.
XX
SQ Sequence 1116 BP; 300 A; 332 C; 230 G; 254 T; 0 other;

Query Match 9.6%; Score 176; DB 21; Length 1116;
Best Local Similarity 53.5%; Pred. No. 1e-45;
Matches 635; Conservative 0; Mismatches 460; Indels 93; Gaps 9;

QY 345 atgaaggtttcttaagatcctctgtcctctatctgcccctgtagaattggaatgtc 404
DB 1 atgaaggttttttttttttttttttttttttttttttttttttttttttttttt 1116
QY 405 tatcagcaacgcgactgtgcagcaatgatgtgacctaaactccgggacaacaggtatg 464
DB 61 tatctacaactgtcgtgtgtgagattc-----aacacaaggaactgtcaggga-- 109
QY 465 gctcagagccgtcgtatgtcgtatcccgacaccccggaacacagcaaatgttga 524
DB 110 -caaccaaacgcatatgtgtcgtgtcctagtgacacagagaataatgcaaatgttga 168
QY 525 agcggttcacagacatatctctctctccgctgcacccacacatctccagtactcg 584
DB 169 tacagttcacagacatatgttccctctactactactacacgtcttcacagtgca 228
QY 585 gggaataactagcgagaatgagactc-----ccatcggtctctggaagaatgta 635
DB 229 ggcaaacactgtgagagagagacactactctccacaacatcactccacagagatgtg 288

```

```

QY 636 ggaacatgcaagatctgcgattatgaagcactgcagcgcacacaggaacacatcaggagc 695
DB 289 ggaacatgcaagatgtgtgtgttaagcatgtgatgcacacaggaacacatcacaacct 348
QY 696 aacacagggtcagggccttgtgaacccacagacagacacacagccttgtgatctcc 755
DB 349 tgcgaaccgggaacagacttggccctctc-----cagccagttagagctatacaat 399
QY 756 accactctcgcgtacccggagactgtgagttctgcaagctccatctgtgtgcaagt 815
DB 400 gccacacactgtgtgtgtctctctgtgcagacgcgcgaagactgttaccatctact 459
QY 816 gtctctgtctgtccacccacagagcagttccaggtccactacacaaatggacatccggt 875
DB 460 ccaaaaggcgtctctgcacagccgacacacatc-----attctcatccaacggcacc 515
QY 876 attctggaagcaacagatccttctgtgcactcttccccaagagcaacttgtgtcag 935
DB 516 aggtactacaatagtgcattcccggaatactactacagtaattgtctctccaaag 575
QY 936 ggaacagccatgcccctgacactccttgacagcaacagatccttctgcaactctcccca 995
DB 576 tgcctctctgtccaagttgtgtccggaacacagagaagctccaaccccgagacacc 635
QY 996 gggtcacttgtgtcagggacagcgccactgacactcctgggacgacaaagttcctt 1055
DB 636 aggcgctgttcacatctggaacacatctccttgccaatatgtacaggaactcaag----- 689
QY 1056 tctgcaactcttccccaagagatcattgtgtcaggaacagcagccctagcaactcct 1115
DB 690 --tagcggaaacaccccttgacagcaatctctgcgggaatcgtctctgttgaagcttc 747
QY 1116 ggagagaacagcgtcctcttgacacactccttccagagagcaacttgtgtcagatcaa 1175
DB 748 cagccggataagctactcttggaacccctacccc----- 782
QY 1176 ggtatgcttggaacatccgagttccttgacagagagagactagtgtgacagtgtgt 1235
DB 783 -----tctgttagcacaagcaactctggaacaga-----atgctgctgc 822
QY 1236 gccctcagatccaacacctgtcatctgcgcacatcatgaacatagtgtgaatgtgtat 1295
DB 823 actctccaatcacaacacctgttaattgcacactctctatgtgttcaacagatgtgtat 882
QY 1296 cctctctcaccgacatcagccaaacctcggatcactgtggaatctgtgtgacatcag 1355
DB 883 cccagctcaactgtcgtacgcccacaccttggaacattgtggactctgcatcgacacacag 942
QY 1356 aagacggggggagacatccttgagacaaaccccgagaagttcgcgcacagtaatgcatt 1415
DB 943 aagtc-----aacatcatccttggaacccaagaagaagcctgttagcacaagtatgaa 996
QY 1416 gaggcctgtgcaacacacaacaggtatattataggaacagagatcattgttga 1475
DB 997 gaaagatgcgtgcacacacactcaacatcgtgttcttaagaaatgtgtgtaactctcctagc 1056
QY 1476 ccaggaatgtacatgaatgaatgaatcctccatgacacactgtgtccaa 1523
DB 1057 ccggggatgtataatcactcaactcccatgcaacgctgtgtgcaa 1104

RESULT 3
AAF45087
ID AAF45087 standard; cDNA: 740 BP.
XX
AC AAF45087;
XX
DT 03-APR-2001 (first entry)
XX
DE Human secreted protein coding sequence SEQ ID NO: 26.
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;

```



```
Db      402 cctgcacacatgggtgtgtgtatccatccacacacatcccttcagccacacggtct 461
Oy      1009 gtccagagacagccagccatagcactctgtggcagacagagtccttccttcgacacttc 1068
Db      462 accacgtccacccctgttcctcctgcacacggtgtgtgtgtatccatccacacacacttc 521
Oy      1069 ccccaaggagtcactctgtctcaggagacagccagcctcctcctctgtggcagacacag 1128
Db      522 cccctcagccacgcgtctctacacagctccacccggttcctcctgcacacgttgggtgtgtaga 581
Oy      1129 ttc 1131
Db      582 tgc 584

RESULT 5
ABA64998
ID      ABA64998 standard; DNA; 611 BP.
XX
AC      ABA64998;
XX
DT      01-FEB-2002 (first entry)
XX
DE      Human foetal liver single exon nucleic acid probe #13303.
XX
KW      Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
OS      Homo sapiens.
XX
PN      WO200157277-A2.
XX
PD      09-AUG-2001.
XX
PF      30-JAN-2001; 2001MO-US00669.
XX
PR      04-FEB-2000; 2000US-0180312.
PR      26-MAY-2000; 2000US-0207456.
PR      30-JUN-2000; 2000US-0608408.
PR      03-AUG-2000; 2000US-0632366.
PR      21-SEP-2000; 2000US-0234687.
PR      27-SEP-2000; 2000US-0236359.
PR      04-OCT-2000; 2000GB-0024263.
XX
PA      (MOLE-) MOLECULAR DYNAMICS INC.
PI      Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR      WPI; 2001-483447/52.
XX
PT      Human genome-derived single exon nucleic acid probes useful for
PT      analyzing gene expression in human fetal liver -
XX
PS      Claim 4; SEQ ID NO 13303; 639pp + sequence listing; English.
XX
CC      The invention relates to a single exon nucleic acid probe for
CC      measuring human gene expression in a sample derived from human foetal
CC      liver. The single exon nucleic acid probes may be used for predicting,
CC      measuring and displaying gene expression in samples derived from human
CC      fetal liver. The present sequence is a single exon nucleic acid
CC      probe of the invention.
CC      Note: The sequence data for this patent did not form part of the
CC      printed specification, but was obtained in electronic format directly
CC      from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SO      Sequence 611 BP; 104 A; 238 C; 111 G; 158 T; 0 other;

Query Match      2.1%; Score 38.2; DB 22; Length 611;
Best Local Similarity 47.3%; Pred. No. 0.21;
Matches 115; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

Oy      889 aacagatccttctgtgcaactctcccccagagcactctgtgtcaggagacgcatatgc 948
```

```
Db      342 accacgtccacccctgttcctcctccttcagccacagtcgtctacacagctcacccgttc 401
Oy      949 ctgacactctctggacacacagatctcttcctgtgacactctcccccagggttactttgt 1008
Db      402 cctgcacacatgggtgtgtgtatccatccacacacatcccttcagccacacggtct 461
Oy      1009 gtccagagacagccacagcctagcactctgtggcagacagagtccttccttcgacacttc 1068
Db      462 accacgtccacccctgttcctcctgcacacggtgtgtgtgtatccatccacacacacttc 521
Oy      1069 ccccaaggagtcactctgtctcaggagacagccagcctcctcctctgtggcagacacag 1128
Db      522 cccctcagccacgcgtctctacacagctccacccggttcctcctgcacacgttgggtgtgtaga 581
Oy      1129 ttc 1131
Db      582 tgc 584

RESULT 6
ABA32106
ID      ABA32106 standard; DNA; 611 BP.
XX
AC      ABA32106;
XX
DT      23-JAN-2002 (first entry)
XX
DE      Probe #10572 for gene expression analysis in human heart cell sample.
XX
KW      Human; gene expression; heart; microarray; vascular system; probe;
KW      cardiovascular disease; hypertension; cardiac arrhythmia;
KW      congenital heart disease; ss.
XX
OS      Homo sapiens.
XX
PN      WO200157274-A2.
XX
PD      09-AUG-2001.
XX
PF      30-JAN-2001; 2001MO-US00666.
XX
PR      04-FEB-2000; 2000US-0180312.
PR      26-MAY-2000; 2000US-0207456.
PR      30-JUN-2000; 2000US-0608408.
PR      03-AUG-2000; 2000US-0632366.
PR      21-SEP-2000; 2000US-0234687.
PR      27-SEP-2000; 2000US-0236359.
PR      04-OCT-2000; 2000GB-0024263.
XX
PA      (MOLE-) MOLECULAR DYNAMICS INC.
PI      Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR      WPI; 2001-488899/53.
XX
PT      Single exon nucleic acid probes for analyzing gene expression in human
PT      hearts -
XX
PS      Claim 4; SEQ ID NO 10572; 530pp; English.
XX
CC      The present invention relates to single exon nucleic acid probes for
CC      measuring human gene expression in a sample derived from human heart. The
CC      present sequence is one such probe. The probes may be used for
CC      predicting, measuring and displaying gene expression in samples derived
CC      from the human heart via microarrays. By measuring gene expression, the
CC      probes are useful for predicting, diagnosing, grading, staging,
CC      monitoring and prognosing diseases of the human heart and vascular system
CC      e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC      congenital heart disease.
CC      Note: The sequence data for this patent did not form part of the printed
CC      specification, but was obtained in electronic format directly from WIPO
CC      at ftp.wipo.int/pub/published_pct_sequences.
```


CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is one of
CC the probes of the invention.

XX Sequence 611 BP; 104 A; 238 C; 111 G; 158 T; 0 other;

Query Match 2.1%; Score 38.2; DB 22; Length 611;

Best Local Similarity 47.3%; Pred. No. 0.21;

Matches 115; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 889 aacagatccttctgacactctcccccagagcacttctgtcaggagagcagcagc 948
DB 342 accacgctccaccctgttccctcgcctccagcagcagctctacacagctcacctgttc 401
QY 949 ctgacactcctgagacagacagatccttctgacactctcccccaggggtcacttgt 1008
DB 402 cctgacacacatgggtgtgtgtgatccatccacacatccctcctcagccacggctct 461
QY 1009 gtccaggacagggcagcagcctcctcctgagcagcagcaggtccttctgacactctc 1068
DB 462 accacgctccaccctgttccctcgcctcagcagcagctgtgtgtgtatccatccacacacttc 521
QY 1069 cccacagagtgacttctgtcaggagcagcagcagcagcagcagcagcagcagcagc 1128
DB 522 cccctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 581
QY 1129 tcc 1131
DB 582 tgc 584

RESULT 9

AA11970
ID AA11970 standard; DNA; 611 BP.

AC AA11970;

DT 12-OCT-2001 (first entry)

DE Probe #9903 for gene expression analysis in human cervical cell sample.

KM Probe: human; microarray; gene expression; cervical epithelial cell;

KW cervical cancer; ss.

XX Homo sapiens.

PN W0200157278-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00670.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-488901/53.

PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -

PS Claim 25; SEQ ID NO 9903; 487bp; English.

XX The present invention relates to human single exon nucleic acid probes
CC (SENPs). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging
CC of diseases of the cervix, notably cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at http://wipo.int/pub/published_pct_sequences.

XX Sequence 611 BP; 104 A; 238 C; 111 G; 158 T; 0 other;

Query Match 2.1%; Score 38.2; DB 22; Length 611;

Best Local Similarity 47.3%; Pred. No. 0.21;

Matches 115; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 889 aacagatccttctgacactctcccccagagcagcacttctgtcaggagagcagcagc 948
DB 342 accacgctccaccctgttccctcgcctccagcagcagctctacacagctcacctgttc 401
QY 949 ctgacactcctgagacagacagatccttctgacactctcccccaggggtcacttgt 1008
DB 402 cctgacacacatgggtgtgtgtgatccatccacacatccctcctcagccacggctct 461
QY 1009 gtccaggacagggcagcagcctcctcctgagcagcagcaggtccttctgacactctc 1068
DB 462 accacgctccaccctgttccctcgcctcagcagcagctgtgtgtgtatccatccacacacttc 521
QY 1069 cccacagagtgacttctgtcaggagcagcagcagcagcagcagcagcagcagcagc 1128
DB 522 cccctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 581
QY 1129 tcc 1131
DB 582 tgc 584

RESULT 10

AA145168
ID AA145168 standard; DNA; 611 BP.

AC AA145168;

DT 17-OCT-2001 (first entry)

DE Probe #13854 used to measure gene expression in human placenta sample.

KM Probe: microarray; human; placenta; antenatal diagnosis;

KW genetic disorder; ss.

XX Homo sapiens.

PN W0200157272-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00663.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI: 2001-488897/53.
DR Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human placenta -
XX
PS Claim 25: SEQ ID No 13854; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SNP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.
XX
SQ Sequence 611 BP; 104 A; 238 C; 111 G; 158 T; 0 other;
XX
Query Match 2.1%; Score 38.2; DB 22; Length 611;
Best Local Similarity 47.3%; Pred. No. 0.21;
Matches 115; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
XX
OY 889 aacagatccttctgacactctcccccagagacacttgtgtcagggagacagcctgc 948
DB 342 accacgtccacccctgttcctcctcagcagcgtgtcctacacgtccaccctgttc 401
OY 949 ctgacactcctgacagcaacagatcctctgacactctcccccagggtacttgt 1008
DB 402 cctgcacacatgggtgtgtgtgtgacatccaccacatccctcctcagccagcgtct 461
OY 1009 gtacaggagacagcagcctagacactcctgtggagacagcaggtccttctgacacttc 1068
DB 462 accacgtccacccctgttcctcctcagcagcgtgtgtgtgtgtgtgtgtgtgtgtgt 521
OY 1069 ccccaagagatcacttgtgtcagggagacagcagcctagcactcctgtggagacagcag 1128
DB 522 cctcagcagccagcgtctacacagctcaccgggttcctcctgacacgtgggtgtgtaga 581
OY 1129 tcc 1131
DB 582 tgc 584
XX
RESULT 11
AAI05680
ID AAI05680 standard; DNA; 611 BP.
XX
AC AAI05680;
XX
DT 09-OCT-2001 (first entry)
XX
DE Probe #5671 used to measure gene expression in human breast sample.
XX
KW Probe; human; breast disease; breast cancer; development disorder; ss;
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS Homo sapiens.
XX
PN WO200157270-A2.
XX
PD 09-AUG-2001.
XX
PF 29-JAN-2001; 2001WO-US00661.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0633665.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI: 2001-476286/51.
XX
XX Novel single exon nucleic acid probe used to measuring gene expression
PT in a human breast -
XX
PS Claim 25: SEQ ID No 5671; 322pp; English.
XX
CC The present invention relates to novel single exon nucleic acid probes.
CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridises at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and
CC non-carcinoma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 611 BP; 104 A; 238 C; 111 G; 158 T; 0 other;
XX
Query Match 2.1%; Score 38.2; DB 22; Length 611;
Best Local Similarity 47.3%; Pred. No. 0.21;
Matches 115; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
XX
OY 889 aacagatccttctgacactctcccccagagacacttgtgtcagggagacagcctgc 948
DB 342 accacgtccacccctgttcctcctcagcagcgtgtcctacacgtccaccctgttc 401
OY 949 ctgacactcctgacagcaacagatccttctgacactctcccccagggtacttgt 1008
DB 402 cctgcacacatgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 461
OY 1009 gtacaggagacagcagcctagacactcctgtggagacagcaggtccttctgacacttc 1068
DB 462 accacgtccacccctgttcctcctcagcagcgtgtgtgtgtgtgtgtgtgtgtgtgt 521
OY 1069 ccccaagagatcacttgtgtcagggagacagcagcctagcactcctgtggagacagcag 1128
DB 522 cctcagcagccagcgtctacacagctcaccgggttcctcctgacacgtgggtgtgtaga 581
OY 1129 tcc 1131
DB 582 tgc 584
XX
RESULT 12
AAO03665
ID AAO03665 standard; DNA; 2465 BP.
XX
AC AAO03665;
XX
DT 07-AUG-1989 (first entry)
XX
DE Sequence homologous to Drosophila Per gene.
XX
KW Sex determination; ruminant embryos; ss.
XX
PN FR2635116-A.
XX
PD 09-FEB-1990.
XX
PR 08-AUG-1988; 88FR-0010706.
PR 08-AUG-1988; 88FR-0010706.
XX
PA (GEOR/) GEORGES M.
XX


```
XX OS Homo sapiens.
XX AC
XX FH Key
XX CDS
XX FT 361..2740
XX FT /tag= a
XX FT /product=MAR subtype ml
XX FT misc_signal
XX FT 282
XX FT /tag= b
XX FT /label=splice acceptor site
XX FT /note=defines 5' end of exon
XX FT 2780
XX FT /tag= c
XX FT polyA_site
XX PN USN7241971-N.
XX PD 14-MAR-1989.
XX PF 08-SEP-1988; 88US-0241971.
XX PR 08-SEP-1988; 88US-0241971.
XX PA (USSH ) NAT INST OF HEALTH.
XX DR WPI; 1989-165452/22.
XX DR P-PSDB; AAR06202.
XX PT Cloned genes for muscarinic acetylcholine receptors -
XX PT for drug screening and diagnostic use.
XX PS Disclosure; ; P; English.
XX SS
XX CC The sequence was indexed from the best available specification copy;
XX CC the unidentifiable bases were annotated as 'N'.
XX CC The sequence may be useful for synthesis of hybridisation probes for
XX CC diagnostic use. The genes are cloned by screening a rat cerebral cortex
XX CC cDNA library with a probe based on nucleotides 170-225 of the non-coding
XX CC strand of porcine brain (ml) cDNA (modified at positions 5, 38 and 53);
XX CC identifying cDNA clones by hybridisation with BamHI or XhoI digests of
XX CC plasmid DNA from 12 cultures contg. up to 500000 independent clones;
XX CC recluturing until a pool contg. less than 5000 indep. clones with a
XX CC single hybridisation band are identified, and isolating individual
XX CC clones by colony hybridisation.
XX CC Stable cell lines produced by transfecting Chinese hamster ovary cells
XX CC (CHO-K1) with various pCD vectors contg. the gene inserts.
XX CC See also AAQ02068-Q020671 and AAQ00086.
XX CC (Note: Revised entry submitted to correct the patent number format of
XX CC US Government-owned NFIS applications to prevent clashes with ongoing US
XX CC granted patent numbers. For further information please visit the Derwent
XX CC web site at www.derwent.com/dwpi/updates/nfis_us.html.)
XX SQ Sequence 2880 BP; 618 A; 921 C; 763 G; 550 T; 28 other;
Query Match 2.0%; Score 37.2; DB 10; Length 2880;
Best Local Similarity 55.4%; Pred. No. 1.2;
Matches 72; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
OY 819 cctgtgtgtcaccaccagcagctcagctactacaacaatggaacatccggtatt 878
DB 1554 cctgtgtgtcaccaccagcagctcagctactacaacaacacatcaccatggtacta 1613
OY 879 cctgtgacagacaacagatccttcttgacactctcccccaggagcaccattgtgtcaggga 938
DB 1614 cgcacctgtgacaacaagccttccgggacacaccttgcctgtctactcttgcgcgcggga 1673
OY 939 caggcctatgc 948
DB 1674 caagagagcgc 1683
```

RESURF 15
AAH99405

```
ID AAH99405 standard; CDNA; 1090 BP.
XX AC
XX FH AAH99405;
XX DT 16-OCT-2001 (first entry)
XX DE Human protein encoding CDNA sequence SEQ ID NO:240.
XX SS
XX CC Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
XX CC antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
XX CC antibacterial; endocrine; cardiac; central nervous system; virucide;
XX CC anti-HIV; fungicide; antimutagen; cardiovascular; antinaemic; anaemia;
XX CC antiagregant; haemostatic; vulnery; antitumor; osteopathic; eczema;
XX CC dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
XX CC neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
XX CC immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
XX CC antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
XX CC cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
XX CC genetic disease; haematopoietic disorder; platelet disorder; asthma;
XX CC thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
XX CC allergic rhinitis; diabetes; multiple sclerosis; depression;
XX CC Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
XX CC neurological disorder; ss.
XX OS Homo sapiens.
XX AC
XX FH WO200153455-A2.
XX PN
XX PD 26-JUL-2001.
XX PF 22-DEC-2000; 2000WO-US35017.
XX PR 23-DEC-1999; 99US-0471275.
XX PR 21-JAN-2000; 2000US-0488725.
XX PR 25-APR-2000; 2000US-0552317.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX DR WPI: 2001-457603/49.
XX DR P-PSDB; AAM25464.
XX SS
XX PT Isolated human polynucleotides encoding polypeptides, useful for the
XX PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
XX PS
XX CC Claim 1: Page 404-405; 1217pp; English.
XX CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
XX CC AAM25963. The proteins can have activities based on the tissues and
XX CC cells they are expressed in, such as: antiinflammatory; antirheumatic;
XX CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;
XX CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
XX CC cardiovascular; antinaemic; antiagregant; haemostatic; vulnery;
XX CC antitumor; osteopathic; dermatological; antiallergic; antiasthmatic;
XX CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
XX CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
XX CC encoding them can be used in gene therapy, antisense therapy and vaccine
XX CC production. The proteins and polynucleotides are useful for screening for
XX CC agonists or antagonists of a protein and for the treatment and diagnosis
XX CC of disorders associated with the activity of a protein e.g. inflammation,
XX CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
XX CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
XX CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
XX CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
XX CC osteoporosis, severe combined immunodeficiency, eczema, allergic
XX CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
XX CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
XX CC neurological disorders.
XX SQ Sequence 1090 BP; 296 A; 182 C; 223 G; 389 T; 0 other;
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 20:05:10 ; Search time 54.88 Seconds
(without alignments)
8190.767 Million cell updates/sec

Title: US-09-755-456-1

Perfect score: 1830
Sequence: 1 gaattcagatgcctcacc.....tcccttagaagtgctgggattc 1830

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :

1: /cgn2_6/prodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/prodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/prodata/2/lna/5A.COMB.seq:*
4: /cgn2_6/prodata/2/lna/5B.COMB.seq:*
5: /cgn2_6/prodata/2/lna/PCITUS.COMB.seq:*
6: /cgn2_6/prodata/2/lna/Backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40.6	2.2	7218	1	US-08-232-463-14
2	34.8	1.9	3448	1	US-08-296-014A-3
3	34.8	1.9	3448	2	US-08-596-405-3
4	34.8	1.9	3448	2	US-08-877-620-3
5	34.8	1.9	4182	1	US-08-296-014A-1
6	34.8	1.9	4182	2	US-08-596-405-1
7	34.8	1.9	4182	2	US-08-877-620-1
8	33.8	1.8	1329	4	US-09-296-284-23
9	33.8	1.8	1337	4	US-09-296-284-2
10	33.8	1.8	4830	4	US-09-296-284-7
11	33.4	1.8	626	4	US-08-943-731-199
12	33.4	1.8	20084	4	US-08-943-731-5
13	33	1.8	1778	4	US-08-934-386-4
14	33	1.8	2661	1	US-08-592-126-70
15	33	1.8	5836	1	US-09-233-086-2
16	31.4	1.7	1672	2	US-09-003-199-1
17	31.2	1.7	1304	2	US-08-766-439-37
18	31.2	1.7	1304	2	US-08-766-439-38
19	31	1.7	1549	2	US-08-865-597A-1
20	30.8	1.7	1370	4	US-09-111-470-9
21	30.8	1.7	1458	4	US-09-111-470-3
22	30.8	1.7	7122	4	US-09-318-448-2
23	30.8	1.7	15297	4	US-09-817-180-3
24	30.6	1.7	1236	3	US-08-643-704A-48
25	30.6	1.7	2339	1	US-08-258-639A-1
26	30.6	1.7	2339	2	US-08-800-951-1
27	30.6	1.7	2339	5	PCT-US95-07391A-1

C 28	30.4	1.7	1212	2	US-09-092-770-18	Sequence 18, Appl
C 29	30.4	1.7	1212	4	US-09-222-851-18	Sequence 18, Appl
C 30	30.4	1.7	4858	3	US-08-436-3328-9	Sequence 9, Appl
C 31	30.2	1.7	4599	1	US-08-431-080-27	Sequence 27, Appl
C 32	30.2	1.7	4599	2	US-08-938-534-27	Sequence 27, Appl
C 33	30.2	1.7	5362	3	US-08-463-210-5	Sequence 5, Appl
C 34	30.2	1.7	12568	2	US-08-387-942C-1	Sequence 11, Appl
C 35	30	1.6	688	4	US-08-998-416-915	Sequence 915, Appl
C 36	30	1.6	2575	4	US-08-858-207A-38	Sequence 38, Appl
C 37	30	1.6	8220	2	US-08-568-459A-11	Sequence 11, Appl
C 38	30	1.6	8220	2	US-08-487-826B-13	Sequence 11, Appl
C 39	30	1.6	19124	2	US-08-487-826B-13	Sequence 13, Appl
C 40	29.8	1.6	1137	3	US-09-082-088-1	Sequence 1, Appl
C 41	29.8	1.6	1137	4	US-09-546-117-1	Sequence 1, Appl
C 42	29.8	1.6	1368	4	US-09-200-673-13	Sequence 13, Appl
C 43	29.8	1.6	1501	1	US-08-349-025-1	Sequence 1, Appl
C 44	29.8	1.6	1501	2	US-08-566-096A-1	Sequence 1, Appl
C 45	29.8	1.6	1501	2	US-08-668-630B-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINER, F.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-F15
US-08-232-463-14

```

Query Match 2.2%; Score 40.6; DB 1 Length 7218;
      Query Similarity 4.7%; Pred. No. 0.028;
Matches 19; Conservative 211; Mismatches 175; Indels 0; Gaps 0

Oy 746 agtgcctccaccactcctgcctcagctgtagctgtgcacagctccatcgt 805
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1026 ATTAAATTCGAGACTTGGCTGGCAGGTCGAGCGAGCTTGCATTTTTTTTTTTTTT 1085

Oy 806 tctgcagatgtctcctgtcgtgtgcctaccacagcagctcagagctactacaatgg 865
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1086 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1145

Oy 866 aaccgcggtactctgacagcagacagatctctctgtcactcttccccagagccac 925
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1146 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1205

Oy 926 ttgtgtcagggacagcgcacatgcctagcactcctctgacagacagatctcttcgac 985
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1206 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1265

Oy 986 tcttccccaggggtcacttgtgtcaggagacagcagcctagcactctggagcaga 1045
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1266 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1325

Oy 1046 acagctctcttcctgacactctccccagagatcacttgtgtcaggagacagccagcc 1105
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1326 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1385

Oy 1106 tagcactcctggcagcagacagctccttcttgacacctctccc 1150
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1386 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1450

RESULT 2
US-08-296-014A-3
: Sequence 3, Application US/08296014A
: Patent No. 5716834
:
: GENERAL INFORMATION:
: APPLICANT: Ding, Jeak Ling
: APPLICANT: Ho, Bow
: TITLE OF INVENTION: The Cloned Factor C cDNA of the
: TITLE OF INVENTION: Singapore Horseshoe Crab, Carcinoscopus
: TITLE OF INVENTION: rotundicauda and Purification of Factor C Proenzyme
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Birch, Stewart, Kolasch & Birch
: STREET: 8110 Gatehouse Road, Suite 500 East
: CITY: Falls Church
: STATE: Virginia
: COUNTRY: USA
: ZIP: 22042
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/296, 014A
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Murphy, Jr., Gerald M.
: REGISTRATION NUMBER: 28,977
: REFERENCE/DOCKET NUMBER: 1781-105P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 205-8000
: TELEFAX: (703) 205-8050
: TELEX: 248345
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3448 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single

```

```

: TOPOLOGY: both
: MOLECULE TYPE: CDNA
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Carcinoscorpius rotundicauda
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 18..3074
: US-08-296-014A-3

Query Match      1 9%: Score 34.8; DB 1; Length 3448;
Best Local Similarity 65.4%; Fred.No.1.2;
Matches 51; Conservative 0; Mismatches 27; Indels 0; Gaps 0.

Oy 1197 gtctctgacacagcagagacgtctagtgcacagtgtgtgccccccagatccaaacct 1256
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1078 gttcttttgacacacctgtgactgtgtgggggtacagccatattaccatgacctttctcagctf 1137
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Oy 1257 gtcatgccgcacatcatg 1274
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1138 gtcgtgacgaccatccatg 1155

RESULT 3
US-08-596-405-3
: Sequence 3, Application US/08596405
: Patent No. 5858706
: GENERAL INFORMATION:
: APPLICANT: Ding, Jeak Ling
: APPLICANT: Ho, Bow
: TITLE OF INVENTION: The Cloned Factor C cDNA of the
: TITLE OF INVENTION: Singapore Horseshoe Crab, Carcinoscorpius
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Birch, Stewart, Kolasch & Birch
: STREET: 8110 Gatehouse Road, Suite 500 East
: CITY: Falls Church
: STATE: Virginia
: COUNTRY: USA
: ZIP: 22042
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/596,405
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Murphy, Jr., Gerald M.
: REGISTRATION NUMBER: 28,977
: REFERENCE/DOCKET NUMBER: 1781-105P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 205-8000
: TELEFAX: (703) 205-8050
: TELEX: 248345
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3448 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: both
: MOLECULE TYPE: CDNA
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Carcinoscorpius rotundicauda
: FEATURE:
: NAME/KEY: CDS
:

```

LOCATION: 18..3074
US-08-596-405-3

Query Match	1.98;	Score 34.8;	DB 2;	Length 3448;
Best Local Similarity	65.48;	Pred. No. 1.2;		
Matches 51;	Conservative 0;	Mismatches 27;	Indels 0;	Gaps 0;

```

QY 1197 gtccctgcacagcagcggaacgctctagtgcgaacagtgtgtgccccctcagaattccaacccct 1256
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1078 GTCCTTGACAGCTGTACTGTGTGGGGTACAGCCATATACATGAACCTTCTCCAGTGT 1137

```

Oy	1257	g	t	c	a	t	g	c	c	g	c	a	t	c	a	t	g	1274
Db	1138	G	T	C	G	T	G	C	A	G	C	A	T	C	C	A	T	1155

RESULT 4
US-08-877-620-3
; Sequence 3, Application US/08877620
; Patent No. 5985590

? APPLICANT: Ding, Jeak Ling
 ? APPLICANT: Ho, Bow
 ? TITLE OF INVENTION: The Cloned Factor C cDNA of the
 ? TITLE OF INVENTION: Singapore Horseshoe Crab, Carinoscopius
 ? TITLE OF INVENTION: rotundicauda and Purification of Factor C Proenzymes
 ? NUMBER OF SEQUENCES: 4
 ? CORRESPONDENCE ADDRESS:

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

1 FILING DATE:
2 CLASSIFICATION:
3 PRIOR APPLICATION DATA:
4 APPLICATION NUMBER: 08/596,405
5 FILING DATE:
6 ATTORNEY/AGENT INFORMATION:
7 NAME: Murphy, Jr., Gerald M.
8 REGISTRATION NUMBER: 28,977
9 REFERENCE/DOCKET NUMBER: 1781-105P
10 TELECOMMUNICATION INFORMATION:
11 TELEPHONE: (703) 205-8000
12 TELEFAX: (703) 205-8050
13 TELEX: 248345
14 INFORMATION FOR SEQ ID NO: 3:

Query Match 1.98; Score 34.8; DB 2; Length 3448;

Best Local Similarity 65.4%; Pred. No. 1.2;
Matches 51; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy 1197 gttcctgacagcaggcgaccgtctcagtggacagtgltgtgccctccagattccaacccct 1256
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1078 GTTCTTTGACAGCGTGCTACTGTGTGGGGCTACAGCCATATACCATGAACCTTTCCTCAGTGT 1137

```
QY 1257 gtcattgcgcattcatg 1274
    ||| ||| ||| ||| ||| |||
Db 1138 GTCGTGACGCATTCATG 1155
```

RESULT 5
US-08-296-014A-1
; Sequence 1, Application US/08296014A

1 APPLICANT: Ding, Jeak Ling
2
3 APPLICANT: Ho, Bow
4
5 TITLE OF INVENTION: The Cloned Factor C cDNA of the
6
7 TITLE OF INVENTION: Singapore Horseshoe Crab, Carcinoscopus
8
9 TITLE OF INVENTION: rotundicauda and Purification of Factor C Proenzym
10
11 NUMBER OF SEQUENCES: 4
12
13 CORRESPONDENCE ADDRESS:
14
15 ADDRESSEE: Birch, Stewart, Kolasch & Birch
16
17 STREET: 8110 Gatehouse Road, Suite 500 East
18
19 CITY: Falls Church
20
21 STATE: Virginia
22
23 COUNTRY: USA

```

1  COMPUTER READABLE FORM:
2  MEDIUM TYPE: Floppy disk
3  COMPUTER: IBM PC compatible
4  OPERATING SYSTEM: PC-DOS/MS-DOS
5  SOFTWARE: PatentIn Release #1.0, Version #1.25
6  CURRENT APPLICATION DATA:

```

```

1 APPLICATION NUMBER: 05/08/296,014A
2
3 FILING DATE:
4
5 CLASSIFICATION: 435
6
7 ATTORNEY/AGENT INFORMATION:
8
9 NAME: Murphy, Jr., Gerald M.
10
11 REGISTRATION NUMBER: 28,977
12
13 REFERENCE/DOCKET NUMBER: 1781-105P
14
15 TELECOMMUNICATION INFORMATION:
16
17 TELEPHONE: (703) 205-8000
18
19 TELEFAX: (703) 205-8050
20
21 TELEX: 248345
22
23 INFORMATION FOR SEQ. ID NO.: 1:
24
25 SEQUENCE CHARACTERISTICS:
26
27 LENGTH: 4182 base pairs
28
29 TYPE: nucleic acid
30
31 STRANDEDNESS: single
32
33 TOPOLOGY: both
34
35 MOLECULE TYPE: cDNA
36
37 HYPOTHEICAL: NO
38
39 ANTI-SENSE: NO
40
41 ORIGINAL SOURCE:
42
43 ORGANISM: Carcinoscopus rotundicauda
44
45 FEATURE:

```

Query Match	1.98	Score	34.8	DB	1	Length	4182
Best Local Similarity	65.48	Pred. No.	1.4				
Matches	51	Conservative	0	Mismatches	27	Indels	0
						Gaps	0

Oy 1197 gtccctgcacagcaggcggaagtcctagtgtgaaagtgttgttggccccctcagaattccaacct 1256
| | | | | | | | | | | | | | | | | | |
Db 1821 gttcttgacagctgctactgtgtgggtmacgcatatatacattgaactttctcagttct 1880

QY 1257 gtcattgcgcattccattg 1274

Db 1881 GTCGTCCAGCCATCCATG 1898
|||||
RESULT 6
US-08-596-405-1
; Sequence 1, Application US/08596405
; Patent No. 5858706
; GENERAL INFORMATION:
; APPLICANT: Ding, Jeak Ling
; APPLICANT: HO, Bow
; TITLE OF INVENTION: The Cloned Factor C cDNA of the
; TITLE OF INVENTION: Singapore Horseshoe Crab, Carinoscorpius
; TITLE OF INVENTION: rotundicauda and Purification of Factor C Proenzyme
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/596,405
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy, Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1781-105P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4182 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Carinoscorpius rotundicauda
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 569..3817
; US-08-596-405-1

Query Match 1.9%; Score 34.8; DB 2; Length 4182;
Best Local Similarity 65.4%; Pred. No. 1.4;
Matches 51; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 1197 gtctctgacagcaggaagctctagtgacagtggtgcccctagattccaacccct 1256
|||||
Db 1821 GTTCTTTGACAGCTGTACTGTGTGGGTACAGCCATATACCATGTAACCTTCTCAGTGT 1880

QY 1257 gtcatgcccgcacatcatg 1274
|||||
Db 1881 GTCGTCCAGCCATCCATG 1898

;; GENERAL INFORMATION:
;; APPLICANT: Ding, Jeak Ling
;; APPLICANT: HO, Bow
;; TITLE OF INVENTION: The Cloned Factor C cDNA of the
;; TITLE OF INVENTION: Singapore Horseshoe Crab, Carinoscorpius
;; TITLE OF INVENTION: rotundicauda and Purification of Factor C Proenzyme
;; NUMBER OF SEQUENCES: 4
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Birch, Stewart, Kolasch & Birch
;; STREET: 8110 Gatehouse Road, Suite 500 East
;; CITY: Falls Church
;; STATE: Virginia
;; COUNTRY: USA
;; ZIP: 22042
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/877,620
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/596,405
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Murphy, Jr., Gerald M.
;; REGISTRATION NUMBER: 28,977
;; REFERENCE/DOCKET NUMBER: 1781-105P
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 205-8000
;; TELEFAX: (703) 205-8050
;; TELEX: 248345
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 4182 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: both
;; MOLECULE TYPE: cDNA
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Carinoscorpius rotundicauda
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 569..3817
;; US-08-877-620-1

Query Match 1.9%; Score 34.8; DB 2; Length 4182;
Best Local Similarity 65.4%; Pred. No. 1.4;
Matches 51; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 1197 gtctctgacagcaggaagctctagtgacagtggtgcccctagattccaacccct 1256
|||||
Db 1821 GTTCTTTGACAGCTGTACTGTGTGGGTACAGCCATATACCATGTAACCTTCTCAGTGT 1880

QY 1257 gtcatgcccgcacatcatg 1274
|||||
Db 1881 GTCGTCCAGCCATCCATG 1898

RESULT 8
US-09-296-284-23
; Sequence 23, Application US/09296284A
; Patent No. 6204040
; GENERAL INFORMATION:
; APPLICANT: Choi, Eun-Sung
; APPLICANT: Rhee, Sang-Ki
; APPLICANT: Lee, Eun-Hae
; TITLE OF INVENTION: Gluconobacter Suboxydans Sorbitol Dehydrogenase, Genes

```

? TITLE OF INVENTION: and Methods of Use Thereof
?
? FILE REFERENCE: 1533_0870000
?
? CURRENT APPLICATION NUMBER: US/09/296,284A
?
? CURRENT FILING DATE: 1999-04-22
?
? NUMBER OF SEQ ID NOS: 87
?
? SOFTWARE: PatentIn Ver. 2.0
?
? SEQ ID NO 23
?
? LENGTH: 1329
?
? TYPE: DNA
?
? ORGANISM: Gluconobacter suboxydans
?
? US-09-296-284-23

```

Query Match	1.8%	Score 33.8	DB 4	Length 1329
Best Local Similarity	50.3%	Pred. NO	1.4	
Matches 83, Conservative	0	Mismatches	82	Indels 0; Gaps 0

QY	1150	caggagaccactctctgccaaggaatcagatagctcggaacaccggagttcctgacagc	1209
Db	1140	ctgggggcaacaacagctcccgggaaacctgtctctgcctcgatalcgcgaagctccgcacatc	1199
QY	1210	agggacagctcagtgagagcagtgctgtgcgccctcagaattccaacctgtcagcgccat	1265
Db	1200	ggggaacgcggttcccaagcgcgcgctggaagctctctccaaggggcttgagtgccataat	1259
QY	1270	ccatgaaacattagtgagaaatcgggtatcccttctctacacgcatata	1314
Db	1260	ggcgacagccttatgtgcgaagagctcgaaactctctcccgacgaacata	1304

RESULT 9
US-09-296-284-2
Sequence 2, Application US/09296284A

; GENERAL INFORMATION:

APPLICANT: Chol, Eun-Sung

APPLICANT: Lee, Eun-Hae

; TITLE OF INVENTION: Gluconobacter Suboxydans Sorbitol Dehydrogenase, Genes
 ; TITLE OF INVENTION: and Methods of Use Thereof

```

; FILE REFERENCE: 1533.0870000
; CURRENT APPLICATION NUMBER: US/09/296,284A

```

```

; CURRENT FILING DATE: 1999-0
; NUMBER OF SEQ ID NOS: 87

```

; SOFTWARE: Pa
; SEQ ID NO 2

LENGTH: 1437
TYPE: DNA

ORGANISM: *Gluconobacter suboxydans*;
DS-09-296-284-2

Query Match	1.88	Score 33.8	DB 4	Length 1437
Best Local Similarity	50.38	Pred. No. 1.5		
Matches 83; Conservative	0	Mismatches 82	Indels 0	Gaps 0

Oy 1150 caggagaccatttctgtcagagatcaacaagtatgccttgaacaatcccgagatctcttgaacgc 1209
 Db 1248 ctggggagaaacaaagctctccgggaacacctgtctgcctctgatactcgaaagctccgcacaa 1307
 Oy 1210 agggacaagctcagtgagacagtgctgtgcccctaaagtttccaacctgtgcacgcgcacat 1265
 Db 1308 gggtaactgcggtttcccaacggccgagcttgaacgtctcttccaagggctgtgattgctcaat 1367
 Oy 1270 ccataagacattatgtgaatatgggtatccctctcttcaacgcgataca 1314
 Db 1368 ggcgcagaccttatgacgaagagctcgaacatctctcccgacgaacaa 1412

RESULT 10
US-09-296-284-7
; Sequence 7, Application US/09296284A
; Patent No. 6204040

```

: GENERAL INFORMATION:
: APPLICANT: Choi, Eun-Sung
: APPLICANT: Rhee, Sang-Ki
: APPLICANT: Lee, Eun-Hae
: TITLE OF INVENTION: Glucanobacter Suboxydans Sobibol Dehydrogenase, Genes
: TITLE OF INVENTION: and Methods of Use Thereof
: FILE REFERENCE: 1533.0870000
: CURRENT APPLICATION NUMBER: US/09/296,284A
: CURRENT FILING DATE: 1999-04-22
: NUMBER OF SEQ. ID NOS: 87
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 7
: LENGTH: 4830
: TYPE: DNA
: ORGANISM: Glucanobacter suboxydans
: US-09-296-284-7

```

Query Match	1.8%	Score 33.8	DB 4	Length 4830
Best Local Similarity	50.3%	Pred No. 3.1		
Matches 83, Conservative	0	Mismatches 82	Indels 0	Gaps 0

QY	1150	caggagccactt	gtgtc	agatca	caagatg	atgc	ctg	gaacatcc	ggaggtc	ctctg	acgc	1209	
Db	4211	ctggggcaac	acgagctc	cgggaa	ccctgt	ctctc	tgga	tatacc	gaagctc	cgcaac	1111	4270	
QY	1210	agggagc	gctctag	gggac	gtgt	gtg	ccccc	ataga	ttc	caaccc	ctgtat	gcggcat	1269
Db	4271	gggactc	gcggttt	cccac	gcgcgc	gcgtga	agctct	cttca	agggc	ctg	atg	gctctacat	4330
QY	1270	ccatgaac	attagt	gaaga	atggg	tatcc	ttctt	cttca	ccg	cataga	1314		
Db	4331	gcgcgag	cttata	tgcg	agagc	ctg	gaac	ctctc	cccg	cgaga	aca	4375	

RESULT 11
US-08-943-731-199/c
; Sequence 199, Application US/08943731

; Patent No. 626515/
; GENERAL INFORMATION:
;

APPLICANT: PROCKOP, DARWIN J.
APPLICANT: SPOTILA, LORETTA D.

;; APPLICANT: DELTAS, CONSTANTINOS D.
; APPLICANT: SEREDA, LARISA

APPLICANT: LARSON, ANDREA W.
APPLICANT: PACK, MICHAEL

; APPLICANT: COLIGE, ALAIN
; APPLICANT: EARLY, JAMES
;

;
; APPLICANT: KORRKO, JARMO
; APPLICANT: ALA-KOKKO, LEENA, et al.

:	TITLE OF INVENTION:	COMPOSITIONS AND METHODS FOR DETECTING
:		
:	TITLE OF INVENTION:	ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES
:		

NUMBER OF SEQUENCES: 666
CORRESPONDENCE ADDRESS:

ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
STREET: ONE COMMERCE SQUARE 2005 MARKET STREET 32ND

STREET: FLR.
CITY: PHILADELPHIA

STATE: PA
COUNTRY: USA

COONIKI: USA
ZIP: 19103-7086
GOVERNMENT BUILDING FOR

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; GENERATED BY: IBM

```

```
;
;
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
```

```

; SOFTWARE: PatentIn Release #1.0,
; CURRENT APPLICATION DATA:

```

;; APPLICATION NUMBER: US/08/943,731
;; FILING DATE: 03-OCT-1997

```

; CLASSIFICATION: 435
;
; PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: US 08/212,322
FILING DATE: 14-MAR-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/803,628
FILING DATE: 03-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: DOYLE LEARY Ph.D., KATHRYN
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: 9598-27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-965-1284
TELEFAX: 215-567-2991
TELEX: 831-494
INFORMATION FOR SEQ ID NO: 199:
SEQUENCE CHARACTERISTICS:
LENGTH: 626 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-943-731-199

Query Match 1.8%; Score 33.4; DB 4; Length 626;
Best Local Similarity 57.0%; Pred No. 1.2;
Matches 61; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 1107 agcactcctgggcaagaaggtccttcttggcacccttcccaaggagcacttgtgt 1166
||| | ||||| ||| | ||| ||| ||| |||
Db 537 ACCTAGTGTGGGAGCGGCGCCACCCGACCAACACCCCTGTGGGTGC 478

Qy 1167 caggatcaaggtatgcttgaacatccggaagtctctggcagcaggg 1213
||| | ||| ||| ||| ||| ||| ||| ||| |||
Db 477 GTGGAGAAACTGTGTGTGATCCACAGATGCTGCAGAGCGACG 431

RESULT 12
US-08-943-731-5/C
Sequence 5, Application US/08943731
Patent No. 6285157

GENERAL INFORMATION:
APPLICANT: PROCKOP, DARWIN J.
APPLICANT: SPOTILA, LORETTA D.
APPLICANT: DELTAS, CONSTANTINOS D.
APPLICANT: SEREDA, LARISSA
APPLICANT: LARSON, ANDREA W.
APPLICANT: PACK, MICHAEL
APPLICANT: COLIGE, ALAIN
APPLICANT: EARLY, JAMES
APPLICANT: KORRKO, JARMO
APPLICANT: ALA-KORRKO, LEENA, et al.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING
TITLE OF INVENTION: ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES
NUMBER OF SEQUENCES: 666
CORRESPONDENCE ADDRESS:
ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND
STREET: FLR.
CITY: PHILADELPHIA
STATE: PA
COUNTRY: USA
ZIP: 19103-7086
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,731
FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,322
FILING DATE: 14-MAR-1994
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/803,628
FILING DATE: 03-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: DOYLE LEARY Ph.D., KATHRYN
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: 9598-27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-965-1284
TELEFAX: 215-567-2991
TELEX: 831-494
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 20084 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-943-731-5

Query Match 1.8%; Score 33.4; DB 4; Length 20084;
Best Local Similarity 57.0%; Pred No. 10;
Matches 61; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 1107 agcactcctgggcaagaaggtccttcttggcacccttcccaaggagcacttgtgt 1166
||| | ||||| ||| | ||| ||| ||| ||| |||
Db 13825 ACCTAGTGTGGGAGCGGCGCCACCCGACCAACACCCCTGTGGGTGC 13766

Qy 1167 caggatcaaggtatgcttgaacatccggaagtctctggcagcaggg 1213
||| | ||| ||| ||| ||| ||| ||| ||| |||
Db 13765 GTGGAGAAACTGTGTGTGATCCACAGATGCTGCAGAGCGACG 13719

RESULT 13
US-08-934-386-4/C

Sequence 4, Application US/08934386
Patent No. 6306636
GENERAL INFORMATION:
APPLICANT: Haselkorn, Robert
APPLICANT: Gornicki, Piotr
TITLE OF INVENTION: Methods for Detecting Nucleic Acid
TITLE OF INVENTION: Segments Encoding Acetyl-CoA Carboxylase
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 433
CITY: Houston
STATE: Texas
COUNTRY: US
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/934,386
FILING DATE: 19-SEP-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: ARSB:521
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1778 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-934-386-4

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 21:33:31 ; Search time 1982.69 Seconds
(without alignments)
12457.534 Million cell updates/sec

Title: US-09-755-456-1

Perfect score: 1830

Sequence: 1 gaattcagatgctcatcac.....tcccttagaagtgtgagatc 1830

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapept 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	43	2.3	720	12	BH535822 BOGEY58TF
2	42.6	2.3	685	12	BH434901 BOHNO30TF
3	41.2	2.3	387	10	BM280783 K105601.Y
4	40	2.2	1000	12	CNS000COQ
5	39.2	2.1	498	10	BM281363
6	39.2	2.1	547	9	AW165709
7	39.2	2.1	565	10	BM280570
8	39.2	2.1	566	10	BM280471
9	39.2	2.1	577	10	BM285287
10	39.2	2.1	614	10	BM284986
11	39.2	2.1	633	10	BM280716
12	39.2	2.1	644	9	AW165781
13	39	2.1	549	10	BM282815
14	39	2.1	699	12	CNS020KA
15	38.6	2.1	311	10	BG955899
16	38.4	2.1	1034	12	CNS03MEV
17	38.2	2.1	405	10	H16015

C 18	38.2	2.1	853	12	CNS03VB2	AL262199 Tetradon
C 19	38	2.1	489	10	BF773988	BF773988 283520 MA
C 20	37.6	2.1	757	12	CNS02JO8	AL200537 Tetradon
C 21	37.6	2.1	803	12	CNS026N7	AL183560 Tetradon
C 22	37.4	2.0	545	12	AQ432826	AQ432826 HS-5132-B
C 23	37.4	2.0	626	10	BI279337	BI279337 UI-R-DAO-
C 24	37	2.0	326	9	AV936678	AV936678 AV936678
C 25	37	2.0	421	10	BI397496	BI397496 AC95 AC-A
C 26	37	2.0	1080	12	CNS05IBJ	AL338680 Tetradon
C 27	36.8	2.0	726	10	BF09183	BF209183 601873016
C 28	36.8	2.0	781	10	BG564009	BG564009 602584842
C 29	36.8	2.0	880	10	BE973122	BE973122 601652003
C 30	36.6	2.0	513	10	BG383812	BG383812 302198 MA
C 31	36.6	2.0	849	12	CNS011T7	AL100693 Drosophila
C 32	36.4	2.0	584	9	AJ392991	AJ392991 AJ392991
C 33	36.4	2.0	899	9	AL537015	AL537015 AL537015
C 34	36.4	2.0	986	9	AL533429	AL533429 AL533429
C 35	36.2	2.0	473	12	AQ116963	AQ116963 HS-3021-A
C 36	36.2	2.0	723	12	CNS01UY4	AL168241 Tetradon
C 37	36.2	2.0	802	12	CNS03VUV	AL262912 Tetradon
C 38	36.2	2.0	1101	12	CNS00L7A	AL068159 Drosophila
C 39	36	2.0	343	9	AI081631	AI081631 ow/77603.5
C 40	36	2.0	677	10	BF703878	BF703878 MI-P-01-a
C 41	36	2.0	685	9	AU132863	AU132863 AU132863
C 42	36	2.0	965	12	CNS01ZDK	AL174161 Tetradon
C 43	36	2.0	1056	10	CNS02URL	AL214842 Tetradon
C 44	35.8	2.0	466	10	BG515679	BG515679 dno03d07.
C 45	35.8	2.0	469	10	BG513854	BG513854 dno03d07.

ALIGNMENTS

RESULT 1
BH535822
LOCUS
DEFINITION BOGEY58TF BOGE Brassica oleracea genomic clone BOGEY58, DNA
ACCESSION BH535822
VERSION BH535822.1 GI:17767517
KEYWORDS
SOURCE
ORGANISM
Brassica oleracea.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 720)
TOWN, C.D., Van Aken, S., Utterback, T. and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other GSSs: BOGEY58TF
COMMENT
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.
location/Qualifiers
1..720
/organism="Brassica oleracea"
/strain="T01000DH3"
/db_xref="taxon:3712"
/clone="BOGEY58"
/clone_11d="BOGE"
/note="Vector: pHOSt; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOSt using BstXI linkers"

FEATURES

BASE COUNT
263 a 100 c 101 g 256 t
ORIGIN

[illegible]

RESULT	2
BH434901/c	
LOCUS	BH434901 685 bp DNA linear GSS 12-DEC-2001
DEFINITION	BOHNQ30TF BOHN Brassica oleracea genomic clone BohnQ30, DNA sequence.
ACCESSION	BH434901
VERSION	BH434901.1 GI:17620622
KEYWORDS	GSS.
SOURCE	Brassica oleracea.
ORGANISM	Brassica oleracea. Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica. 1 (bases 1 to 685) Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M. Whole genome shotgun sequencing of Brassica oleracea unpublished (2001) Other_GSSs: BOHNQ30TR Contact: Chris Town TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA. Tel: 301-838-3523 Fax: 301-838-0208 Email: cdtown@tigr.org DNA is from a doubled haploid provided by Tom Osborn. Seq primer: TF Class: sheared ends.
FEATURES	location/qualifiers
Source	1..685 /organism="Brassica oleracea" /strain="TO1000DH3" /db_xref="taxon:3712" /clone="BOHNQ30" /clone_lib="BOHN" /note="vector: pHOSt1; site_1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOSt1 using BstXI linkers"
BASE COUNT	189 a 105 c 144 g 24 t
ORIGIN	

[illegible]

QY	419	actgt	423
Db	504	AGTGT	500

RESULT	3
BM280783	
LOCUS	387 bp mRNA linear EST 28-DEC-2001
DEFINITION	k105a01.y1 Ascaris suum LA PSORX1 Zarlenga v1 Ascaris suum cDNA 5'
	similar to wp:CE25434 Y54E10Bc.2 collagen ; , mRNA sequence.
ACCESSION	BM280783
VERSION	BM280783.1 GI:17989825
KEYWORDS	EST.
SOURCE	pig roundworm.
ORGANISM	Ascaris suum

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 387)	McCarter, J., Clifton, S., Chapelin, B., Pape, D., Martin, J., Wylie, T., Dante, M., Marra, M., Hiller, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C., Tsagaris, R., Ronto, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T., Harvey, N., Schurr, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.	The Washington Univ. Nematode EST Project, 1999	Unpublished (1999)	Contact: McCarter JP

The library was supplied by Dr. Dante Zarlenga of the USDA Immunology and Disease Resistance Lab in Beltsville, MD. The cDNAs were made from 21 day L4 proximal (isolated from jejunum) and directionally cloned into the NotI/SalI sites of pSPORTL. Seg direction: -40RP from Gibco. High quality sequence stop: 386.

FEATURES	SOURCE
Location/Qualifiers	1. .387
/organism="Ascaris suum"	
/db_xref="taxon:6253"	
/clone_lib="Ascaris suum L4 pSPOR1 zarlenga v1"	
/dev_stage="L4"	
/lab_host="DH5-alpha"	
/note="Vector: pSPOR1 (Life Technologies); Site_1: NotI; Site_2: SalI; The library was supplied by Dr. Dante Zarlenga of the USDA Immunology and Disease Resistance Lab in Beltsville, MD. The cDNAs were made from 21 day L4 proximal (isolated from jejunum) and directionally cloned into the NotI/SalI sites of pSPOR1."	
89 a	125 c 110 g 63 t

Query Match	2.3%	Score 41.2	DB 10	Length 387
Best Local Similarity	50.5%	Pred. No. 1.3		
Matches 100	Conservative	0	Mismatches 98	Indels 0
				Gaps 0
Qy 1048	aggtcctcttcggcacccttcctcccgagagcatttgctcagggaaagccacgccta	1107		
Db 141	ATGACCTCTCCGGGTCCACCAGGCAAAAGTGGTGTCTCCCGGTGATGAGACTTGCCCGCCAG	200		
Qy 1108	gcactcctcgggcagcgaacaggtcctcttccttcggcacccctctccaggaagcacaattgctc	1167		
Db 201	GACCTCCAGGACCAAGATCCGAACTTTCAGATGCGCTCTTGCCGTACACCAACATTC	260		
Qy 1168	agagatcaagatgatgcctcctggaacatccggagttcctctgacagcagagagcagctcgttgac	1227		
Db 261	CTGTCAAGGCTCTCTCTGGACCTCTCTGGGACCTCTCTGGACCACTGGACAGATGGACAC	320		

RESULT	4	1000 bp	DNA	linear	GSS 04-JUN-1999
CONS00COO/c	CNS00COO	1000 bp	DNA	linear	GSS 04-JUN-1999
LOCUS	Drosophila melanogaster genome survey sequence TET3 end of BAC #				
DEFINITION	BACR26P04 of RPCR-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL059446				
VERSION	AL059446.1				
KEYWORDS	GSS.				
SOURCE	fruit fly.				
ORGANISM	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 1000)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoss in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCR-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp. the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://BACPAC.med.buffalo.edu/drosophila_bac.htm .				
FEATURES	Location/Qualifiers				
source	1..1000				
	/organism="Drosophila melanogaster"				
	/db_xref="taxon:7227"				
	/clone_lib="RPCI-98"				
	/clone="BACR26P04"				
	/note="end : TET3"				
BASE COUNT	75 a 181 c 63 g 153 t 528 others				
ORIGIN					
Query Match	2.2%; Score 40; DB 12; Length 1000;				
Best Local Similarity	14.7%; Pred. No. 4;				
Matches	50; Conservative 139; Mismatches 150; Indels 0; Gaps 0;				
Y	1309 catagagcccaaacctcgatcacctggtatctctgtgtgacatacagaagaagggggga	1368			
Db	997 CMTATMMATMTVYANNHCYATWCTTTHHTCTYGGGDDNNKCGKMWYGASKRRKDCRCYGG	938			
Y	1369 catcttcgagcaaaacccggaagatctcgacacgagatctcgaatgagcgtgtgcaa	1428			
Db	937 NRKSGNPTGCAVARNHGHNDVAVNVAABACIMNVAADMDMGHGKAKABACDCKASRYAB	878			
Y	1429 caccacaaccacaggtattataggcaaacagcgatctctgttggaccaggaatgtaca	1488			
Db	877 SHGMSABRCGCBVRYVAGVNDANTVTANAHHMAAKNAHNNNTAAWMAVAAHVAM	818			
Y	1489 atgcataactctcctcgaacactgtctgtccaatgctcgtcaggctaaataaagaag	1548			
Db	817 ABMMAMVYDMCKCKCYVMNNAMVYHNBACMKCTGKWRBAGMDNKANAAAGATYRCGACVY	758			
Y	1549 tttaactctcttctctcgtctcttttggaaacttgatbtggagatgagagagatctatgg	1608			

Db	757	NRASAAAVKNNMTKTKNNKNNKNTKTNTKKKTDJDKTGKKKKTTGKKKMAADAMPDWDAPAR	698
Qy	1609	ctgaagtgaatgccaacactctctctgcccagaagacac	1647
Db	697	RAAMAAYRGDDKRGSTRDASGHMBMYMTMCHSNABHMK	659
RESULT	5		
BM281363			
LOCUS			
DEFINITION	BM281363	498 bp	linear
ACCESSION	BM281363		EST
VERSION	BM281363.1	GI:17990405	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
SOURCE			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
Qy	1048	aggctctctctgacactcttccccagaagtcacttctgtctgaaggagacagccacta	1107
Db	283	ATGGACCTCCGGGTGACACCAAGATGGTGCCTCCGGTATGATGATGACCTGGCCGCGCAG	342
Qy	1108	gcacccctggggagagaacagtcctctctctgacaccccttctccagagagcaacttggtc	1167
Db	343	GACCTCCAGGACCAATGCCGAATCTTCACGATGCTCTTCTTGCCCGTACCAACATATGC	402
Qy	1168	aggatcaggtatgctctggaacatccgagatctctctggaacagcagggacagctctagtgc	1227

Db 403 CCGTCAACCTCCTCGGACCTCTGGGCTTCCTGGACCACTGGACAGATGACAAC 462

RESULT 6

LOCUS AM165709

DEFINITION MBABMA022M13R Ascaris suum (parasitic nematode) body wall muscle and hypodermis Ascaris suum cDNA clone MBABMA022 5', mRNA sequence.

ACCESSION AM165709

VERSION AM165709

KEYWORDS EST.

SOURCE pig roundworm.

ORGANISM Ascaris suum

REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Ascarididae; Ascaridoidea

AUTHORS 1 (bases 1 to 547)

TITLE Daud,J., Geary,T. and Blaxter,M.

JOURNAL A survey of genes expressed in the parasitic nematode Ascaris suum

COMMENT Unpublished (2000)

Contact: Blaxter ML

Institute of Cell, Animal and Population Biology

University of Edinburgh

Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9

3JT, UK.

Tel: +44 131 650 6760

Fax: +44 131 670 5450

Email: mark.blaxter@ed.ac.uk

The Ascaris suum EST dataset (including the ASC clustering information) is available on the www at http://www.ed.ac.uk/tlide

(mbx/Ascarisweb/AscarisEST.html)

PCR Primers

FORWARD: M13 Reverse (AGCGGATTAACAATTTCACACAGA)

BACKWARD: M13 Forward (CGCGGATTTTCCACGACAGC)

Seq primer: M13 Reverse (AGCGGATTAACAATTTCACACAGA).

FEATURES

Source

Location/Qualifiers

1..547

/organism="Ascaris suum"

/db_xref="taxon:6253"

/clone.lib="MBABMA022"

/clone.lib="Ascaris suum (parasitic nematode) body wall muscle and hypodermis"

/sex="mixed"

/tissue_type="body wall muscle and hypodermis"

/dev_stage="adult"

/note="Vector: Lambda Zap II; Site_1: NotI (5'end); Site_2: NotI (3'end); Ascaris suum is an intestinal nematode parasite of pigs. The library was constructed from dissected body wall muscle and hypodermis tissue for Dr. T. Geary, Pharmacia-Upjohn Inc, Kalamazoo, MI, USA [tgeary@am.pnu.com]"

BASE COUNT 121 a 172 c 152 g 102 t

ORIGIN

Query Match

Best Local Similarity 51.1%; Score 39.2; DB 9; Length 547;

Matches 92; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 1048 aggtccttctggaactctcccccagagtgtaactttgtcagagagagccagccta 1107

DB 242 ATGGAGCTCCGGGTGCACAGCAAGATGATGTCCTCCGCTGATGAGCACTGGCCCGCAG 301

QY 1108 gaactccttgagcaacaagtgctcttcgcaacctctccagaggaactttgtc 1167

DB 302 GACCTCCAGAGCAAGATGCCGAATTCACAGATCGTCTTGCCCGGACACCAACATGTC 361

QY 1168 aggatcaagtagtgcctgtaacatccgaggtctcctgagagagaggaactagtgac 1227

DB 362 CCGTCAACCTCCTCGGACCTCTGGGCTTCCTGGACCACTGGACAGATGACAAC 421

RESULT 7

BM280570

LOCUS BM280570

DEFINITION K102d10.y1 Ascaris suum L4 pSPORT1 Zarlenga v1 Ascaris suum cDNA 5' similar to WP:CE25434 Y54E10BL.2 collagen ;, mRNA sequence.

ACCESSION BM280570

VERSION BM280570

KEYWORDS EST.

SOURCE pig roundworm.

ORGANISM Ascaris suum

REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Ascarididae; Ascaridoidea

AUTHORS 1 (bases 1 to 565)

McCartier,D., Clifton,S., Chlapelli,B., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagarelisvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe,M., Allen,M., Pearson,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.

The Washington Univ. Nematode EST Project, 1999

Unpublished (1999)

Contact: McCarter JP

The Washington Univ. Nematode EST Project, 1999

Washington University School of Medicine

444 Forest Park Parkway, Box 850L, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

The library was supplied by Dr. Dante Zarlenga of the USDA Immunology and Disease Resistance Lab in Beltsville, MD. The cDNAs were made from 21 day L4 proximal (isolated from jejunum) and directionally cloned into the NotI/SalI sites of pSPORT1.

Seq primer: -40RP from Gibco

High quality sequence stop: 481.

FEATURES

Source

Location/Qualifiers

1..565

/organism="Ascaris suum"

/db_xref="taxon:6253"

/clone.lib="Ascaris suum L4 pSPORT1 Zarlenga v1"

/dev_stage="L4"

/lab_host="DH5-alpha"

/note="Vector: pSPORT1 (Life Technologies); Site_1: NotI; Site_2: SalI; The library was supplied by Dr. Dante Zarlenga of the USDA Immunology and Disease Resistance Lab in Beltsville, MD. The cDNAs were made from 21 day L4 proximal (isolated from jejunum) and directionally cloned into the NotI/SalI sites of pSPORT1."

BASE COUNT 118 a 176 c 157 g 114 t

ORIGIN

Query Match

Best Local Similarity 51.1%; Score 39.2; DB 10; Length 565;

Matches 92; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 1048 aggtccttctggaactctcccccagagtgtaactttgtcagagagagccagccta 1107

DB 270 ATGGAGCTCCGGGTGCACAGCAAGATGATGTCCTCCGCTGATGAGCACTGGCCCGCAG 329

QY 1108 gaactccttgagcaacaagtgctcttcgcaacctctccagaggaactttgtc 1167

DB 330 GACCTCCAGAGCAAGATGCCGAATTCACAGATCGTCTTGCCCGGACACCAATGTC 389

QY 1168 aggatcaagtagtgcctgtaacatccgaggtctcctgagagagaggaactagtgac 1227

DB 390 CCGTCAACCTCCTCGGACCTCTGGGCTTCCTGGACCACTGGACAGATGACAAC 449

RESULT 8

BM280471

LOCUS BM280471

DEFINITION K101b10.y1 Ascaris suum L4 pSPORT1 Zarlenga v1 Ascaris suum cDNA 5' similar to WP:CE25434 Y54E10BL.2 collagen ;, mRNA sequence.

[illegible]

SOURCE	ORGANISM	plg roundworm.	Ascaris suum	Eukaryota: Metazoa: Nematoda; Chromadorea; Ascaridida; Ascaridoidea
REFERENCE	AUTHORS	McCarter, J., Clifton, S., Chappell, B., Pape, D., Martin, J., Wylie, T., (bases 1 to 577)	1 (bases 1 to 577)	
TITLE	JOURNAL	The Washington Univ. Nematode EST Project, 1999	Unpublished (1999)	
COMMENT	CONTACT	McCarter JP	The Washington Univ. Nematode EST Project, 1999	
		Washington University School of Medicine	444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA	
		Tel: 314 286 1800		
		Fax: 314 286 1810		
FEATURES	SOURCE	Email: est@wustl.wustl.edu		
		The library was supplied by Dr. Dante Zarlenga of the USDA		
		Immunology and Disease Resistance Lab in Beltsville, MD. The cDNAs		
		were made from 21 day L4 proximal (isolated from jejunum) and		
		directionally cloned into the NotI/SalI sites of pSPOR1.		
		Seq primer: -40RP from Glbco		
		High quality sequence stop: 499.		
		Location/Qualifiers		
		1..577		
		/organism="Ascaris suum"		
		/db_xref="taxon:6253"		
		/clone_11b="Ascaris suum L4 pSPOR1 Zarlenga v1"		
		/dev_stage="L4"		
		/lab_host="DH5-alpha"		
		/note="Vector: pSPOR1 (Life Technologies); Site_1: NotI;		
		Site_2: SalI; The library was supplied by Dr. Dante		
		Zarlenga of the USDA Immunology and Disease Resistance Lab		
		in Beltsville, MD. The cDNAs were made from 21 day L4		
		proximal (isolated from jejunum) and directionally cloned		
		into the NotI/SalI sites of pSPOR1."		
BASE COUNT		122 a 182 c 156 g 117 t		
ORIGIN				
Query Match		2.1%; Score 39.2; DB 10; Length 577;		
Best Local Similarity		51.1%; Pred. No.5.7;		
Matches		92; Conservative 0; Mismatches 88; Indels 0; Gaps 0;		
QY	1048	aggctccttctgacactcttccccagagatcattgtgtcagagacagccagccta	1107	
DB	289	ATGGACCTCCGGGTGACACGCAAGATGGTCCGCGTGTGAGACTGCGCCGCG	348	
QY	1108	gcattccctggagagaacagagtccttctctgacacctctcccaaggagccattgttc	1167	
DB	349	GACCTCCAGGACCAATGCGCAATCTTCAGATGCTCTTCCCGCTACCAACATGTC	408	
QY	1168	aggaacagatgatgctggaacatccggagatctcgagacagagagacagcttagtgag	1227	
DB	409	CCTGTCAAGCTCTCTGTGACTCTCTGGCGCTCTGTGACCACTGTGACAGATGACAC	468	
RESULT	10			
LOCUS	BM284986	614 bp	mrna	linear
DEFINITION	hm3d05.y1 Ascaris suum L4 pSPOR1 Zarlenga v1 Ascaris suum cDNA 5			
	similar to wp:CE35434 Y54E10B1.2 collagen, contains MER22.b1 MSRI			
	repetitive element ; mRNA sequence.			
ACCESSION	BM284986			
VERSION	BM284986.1	GI:17994028		
KEYWORDS	EST.			
SOURCE	plg roundworm.			
ORGANISM	Ascaris suum			
	Eukaryota: Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea			

REFERENCE
AUTHORS

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 614)	Ascarididae: Ascaris.			
	McCarter, J., Clifton, S., Chispebell, B., Page, D., Martin, J., Wylie, T., Dante, M., Marra, M., Hillier, L., Kucab, T., Theisner, B., Bowers, Y., Gibbons, M., Rilter, E., Bennett, J., Franklin, C., Tsagarisvilli, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K., Steple, M., Allen, M., Person, B., Swaller, T., Harvey, N., Schurk, R., Kohn, S., Shilt, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.			
	The Washington Univ. Nematode EST Project, 1999			
	Unpublished (1999)			
	Contact: McCarter JP			

TITLE
JOURNAL
COMMENT

The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)
Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
The library was supplied by Dr. Dante Zarlena of the USDA
Immunology and Disease Resistance Lab in Beltsville, MD. The cDNAs
were made from 21 day L4 proximal (isolated from jejunum) and
directionally cloned into the NotI/SalI sites of pSPOR1.
Seq primer: -40RP from Glbco
High quality sequence stop: 483.

FEATURES

```

129 a 196 c 170 g 119 t
    /organism="Ascaris suum"
    /db_xref="taxon:6253"
    /clone_lib="Ascaris suum I4 pSPOR1 zarlenga v1"
    /dev_stage="I4"
    /lab_host="DH5-alpha"
    /note="Vector: pSPOR1 (Life Technologies); Site_1: NotI;
    Site_2: SalI; The library was supplied by Dr. Dante
    Zarlenga of the USDA Immunology and Disease Resistance Lab
    in Beltsville, MD. The cDNAs were made from 21 day I4
    proximal (isolated from jejunum) and directionally cloned
    into the NotI/SalI sites of pSPOR1."

```

BASE COUNT
ORIGIN

Query Match	Score	DB	Length
2.18;	39.2;	10;	614;

Best Local Similarity	51.16;	Pred.NO. 3.8;							
Matches	92;	Conservative	0;	Mismatches	88;	Indels	0;	Gaps	0;

1048 aggtccttctgacactcttccccagagagtcacttgtgtcagggacagccacgccta 1107

Db 268 ATGACCTCCGGGTGCACCAGGCAAGATGTTGCTCCCGGTGATGGACCTGGCCCCGAC 327

QY 1108 gcaactcctgggcagcaacaggtctcttctgcaacctctctccagagaccacttgtgtc 1167

Db 328 GACCTCCAGGACCAGATGCCGAAGTTCACGATCGTCTTCTGCCCCGTACCAACACAATGTC 387

QY 1168 aggatcaaggtatgccttgaacatccgaggttccttgacagcgagcagtcctagtgc 1227

Db 388 CCCTGTCAGCTCCTCCTGGACCTCCTGGCCCTCCTGGACCACTGGACAGGATGGACAAC 447

RESULT 1

Accession	Definition	LOCUS
B0280716		
B0280716	633 bp mRNA linear EST 28-DEC-2001	
K104b10.v1	Ascaris suum I4 pSOPRI Zarlena v1 Ascaris suum cDNA 5 similar to wv:CM25434 Y54E10B1.2 collagen ; mRNA sequence.	
B0280716		

VERSION
REVISED

SOURCE
ORGANISMS

Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea
: Ascarididae; Ascaris

REFERENCE

AUTHORS

, Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B., Bowers, Y.

TITLE
JOURNAL
COMMENT

Gibbons, R., Ritter, E., Bennett, J., Franklin, C., Tsagaratshilli, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K., Steple, M., Allen, M., Person, B., Swaller, T., Harvey, N., Schurk, R., Kohn, S., Shu, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., Wilson, R.

The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)

Contact: McCarter, JP

The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu

The library was supplied by Dr. Dante Zarlinga of the USDA Immunology and Disease Resistance Lab in Beltsville, MD. The cDNAs were made from 21 day I4 proximal (isolated from jejunum) and directionally cloned into the NotI/SalI sites of pSPORT1.

Seq primer: -40RP from Gibco

High quality sequence stop: 501.

FEATURES
Source

```

/organism="Ascaris suum"
/db_xref="taxon:6253"
/clone.lib="Ascaris suum L4 pSPOR1 zarlenga v1"
/dev_stage="L4"
/lab_host="DH5-alpha"
/notes="Vector: pSPOR1 (Life Technologies); Site_1: NotI;
Site_2: SalI; The library was supplied by Dr. Dante
Zarlenga of the USDA Immunology and Disease Resistance Lab
in Beltsville, MD. The cDNAs were made from 21 day L4
proximal (isolated from jejunum) and directionally cloned
into the NotI/SalI sites of pSPOR1."

```

BASE COUNT
ORIGIN

Query Match	2.1%;	Score 39.2;	DB 10;	Length 633;
Best Local Similarity	51.1%;	Pred. No. 5.9;		
Matches 92;	Conservative	0;	Mismatches 88;	Indels 0;
				Gaps 0

QY 1048 aggtccttctgcaactctccccagagagtcacttgtgtcagggacagccacgccta 1107

Db 289 ATGACCTCCGGGTGCACCAAGCAAGATGTTCTCCCGGTGATGGACCTGCCCCGCCAG 348

QY 1108 gcactcctgggcagcaacaggtccttcttgcaacctctccagagagccacttgtgc 1167

Db 349 GACCTCAGACCAGATGCCGAATTCACGATGTTCTTCTGCCCGTACCACCACAATGTC 408

QY 1168 aggatcaaggtatgcctggaacatccggaattcctgacagcaggacagtctagtggac 1227

Db 409 CCTGTCAAGCTCCTCCTGGACCTCCTGGGCTCCTGGACCACTGGACAGATGGACAAC 468

RESULT 13

AW165781
LOCUS

LOCUS	DEFINITION
AM165781	644 bp mRNA linear EST 12-NOV-1998
MBASBAM197M13R	Ascaris suum (parasitic nematode) body wall muscle and hypodermis
Ascaris suum cDNA clone MBASBAM197 5', mRNA sequence.	

VERSION

SOURCE

REFERENCE
AUTHORS

TITLE
JOURNAL

COMMENT

A survey of genes expressed in the parasitic nematode *Ascaris suum*
Unpublished (2000)
Contact: Blaxter ML
Institute of Cell, Animal and Population Biology

University of Edinburgh
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
3JT, UK.
Tel: +44 131 650 6760
Fax: +44 131 670 5450
Email: mark.blaxter@ed.ac.uk
The Ascaris suum EST dataset (including the ASC clustering
information) is available on the www at <http://www.ed.ac.uk/~tilde/mblx/AscarisWeb/AscarisEST.html>
PCR Primers
FORWARD: M13 Reverse (AGCGATTAACATTTCCACACAGA)
BACKWARD: M13 Forward (CGCCAGGCTTCCGAGTCACAGC)
Seq primer: M13 Reverse (AGCGATTAACATTTCCACACAGA).
Location/Qualifiers
1. .644

FEATURES

source
/organism="Ascaris suum"
/db_xref="taxon:6253"
/clone="MBA8BMA197"
/clone_1lb="Ascaris suum (parasitic nematode) body wall
muscle and hypodermis"
/sex="mixed"
/tissue.type="body wall muscle and hypodermis"
/dev_stage="adult"
/note="Vector: lambda zap II; Site_1: NotI (5'end);
Site_2: NotI (3'end); Ascaris suum is an intestinal
nematode parasite of pigs. The library was constructed
from dissected body wall muscle and hypodermis tissue for
Dr. T. Geary, Pharmacla-Upjohn Inc, Kalamazoo, MI, USA
[tgeary@em.pnu.com]"
BASE COUNT 144 a 205 c 179 g 115 t 1 others
ORIGIN

Query Match 2.1%; Score 39.2; DB 9; Length 644;
Best Local Similarity 51.1%; Pred. No. 5.9;
Matches 92; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 1048 aggtccttcctgacactctcccccagagatctgtgtcagggagcagccacta 1107
DB 228 ATGACCTCCGGGACACAGCAAGATGCTCCGGGAGAGACCTGCCGCCAG 287
QY 1108 gcaatctggagcagcaagctcttcctgacactctcccccagggagcactgtgtc 1167
DB 288 GACCTCCAGGACGAGATGCCAAGCTTCAGATCTCTTCCGCCGTACACACATGTC 347
QY 1168 agatcaagatgctcgtgaacatccgagctcgtgagcagcagagctagtgtac 1227
DB 348 CCTGTCAAGCTCCTCTGACCTCTGCGCTCTGTGACACCTGACAGATGACAC 407

RESULT 13 549 bp mRNA linear EST 19-JUL-2001
BI282815 LOCUS
DEFINITION UT-R-DA0-byp-e-09-0-UI-R-DA0 Rattus norvegicus cDNA clone
ACCESSION BI282815
VERSION BI282815.1 GI:14933954
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 549)
AUTHORS Bonaldi, M.F., Lennon, G. and Soares, M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to verify it as a clone from the
non-normalized rat eye library cDNA library preparation: M.B.
Soares Lab Clone distribution: clones will be available through
Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA=yes.

FEATURES

source
Location/Qualifiers
1. .549
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-DA0-byp-e-09-0-UI"
/clone_1lb="UI-R-DA0"
/dev_stage="ADULT"
/lab_host="DH10B (Life Technologies)"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-DA0
library is a non-normalized library constructed from rat
salivary gland tissue. For a detailed description of the
library from which this clone was derived, please visit
our web site at ratat.eng.uiowa.edu. The subtraction has
been previously described in (Bonaldi, Lennon and Soares,
Genome Research 6:791-806, 1996)
TAG_1lb=UI-R-DA0
TAG_TISSUE=rat eye
TAG_SEQ=CAGCC"
BASE COUNT 55 a 118 c 201 g 175 t
ORIGIN

Query Match 2.1%; Score 39; DB 10; Length 549;
Best Local Similarity 46.2%; Pred. No. 6.4;
Matches 129; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 882 ggaacagcaacagatctcttcgtgacactctcccccagagcactgtgtcagggagac 941
DB 496 GGCCACGACGAGACCTCAAGGCCACCCACACAGAGAGACCAAGAGACTCA 437
QY 942 gcatgcttagcatctccgtgacagcaagatctcttcgtgacactctcccccagggttc 1001
DB 436 GGCCACCCCAACCAAGAGAGACCAAGACTCTAGGGCCCAACCCCAAGAGAGGC 377
QY 1002 actgtgtcagggagcagcgcctgacactcctgtgagcagcaaggtctcttcgtgc 1061
DB 376 CCACAGCAGAGACCTCAAGGCCCAACCCCAAGAGAGACCAAGAGACTCAAGGC 317
QY 1062 actctcccccagagtaactgtgtcaggaagcagcagccttagcactctcgtggagc 1121
DB 316 CCACCCCAACCAAGAGAGACCAAGAGACTCTAGGGCCCAACCCCAAGAGAGCA 257
QY 1122 caacagctcttcgtgacactctcctccagagcact 1160
DB 256 CAGCAGAGACTCAAGGCCCAACCCCAAGAGTGGCCT 218

RESULT 14 699 bp DNA linear GSS 14-MAY-2000
CNS020KA LOCUS
DEFINITION Tetradon nigroviridis genome survey sequence PUC-ori end of clone
153D20 of library G from Tetradon nigroviridis, genomic survey
sequence.
ACCESSION AL206803
VERSION AL206803.1 GI:7865622
KEYWORDS GSS; genome survey sequence.
SOURCE Tetradon nigroviridis
ORGANISM Tetradon nigroviridis

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2002, 00:45:01 ; Search time 3374.16 Seconds

(without alignments)
11733.843 Million cell updates/sec

Title: US-09-755-456-1

Perfect score: 1830
1 gaattcagatgcctcatacc.....tcccttaagaatgtyggatc 1830

Sequence: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Scoring table: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

```

1: /cgn2_6/ptodata/2/pna/PCTUS.COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06.COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07.COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US08.COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US082.COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US083.COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US084.COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US085.COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US086.COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US087.COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US088.COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US089.COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US090.COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US091.COMB.seq.*
16: /cgn2_6/ptodata/2/pna/US092.COMB.seq.*
17: /cgn2_6/ptodata/2/pna/US093.COMB.seq.*
18: /cgn2_6/ptodata/2/pna/US094.COMB.seq.*
19: /cgn2_6/ptodata/2/pna/US095A.COMB.seq.*
20: /cgn2_6/ptodata/2/pna/US095B.COMB.seq.*
21: /cgn2_6/ptodata/2/pna/US095C.COMB.seq.*
22: /cgn2_6/ptodata/2/pna/US096.COMB.seq.*
23: /cgn2_6/ptodata/2/pna/US096A.COMB.seq.*
24: /cgn2_6/ptodata/2/pna/US096B.COMB.seq.*
25: /cgn2_6/ptodata/2/pna/US096C.COMB.seq.*
26: /cgn2_6/ptodata/2/pna/US096D.COMB.seq.*
27: /cgn2_6/ptodata/2/pna/US096E.COMB.seq.*
28: /cgn2_6/ptodata/2/pna/US097A.COMB.seq.*
29: /cgn2_6/ptodata/2/pna/US097B.COMB.seq.*
30: /cgn2_6/ptodata/2/pna/US097C.COMB.seq.*
31: /cgn2_6/ptodata/2/pna/US098A.COMB.seq.*
32: /cgn2_6/ptodata/2/pna/US098B.COMB.seq.*
33: /cgn2_6/ptodata/2/pna/US098C.COMB.seq.*
34: /cgn2_6/ptodata/2/pna/US099A.COMB.seq.*
35: /cgn2_6/ptodata/2/pna/US099B.COMB.seq.*
36: /cgn2_6/ptodata/2/pna/US099C.COMB.seq.*
37: /cgn2_6/ptodata/2/pna/US100.COMB.seq.*
38: /cgn2_6/ptodata/2/pna/US101.COMB.seq.*
39: /cgn2_6/ptodata/2/pna/US6000.COMB.seq.*
40: /cgn2_6/ptodata/2/pna/US6001.COMB.seq.*
41: /cgn2_6/ptodata/2/pna/US6002.COMB.seq.*
42: /cgn2_6/ptodata/2/pna/US6003.COMB.seq.*
43: /cgn2_6/ptodata/2/pna/US6004.COMB.seq.*

```

```

44: /cgn2_6/ptodata/2/pna/US6005.COMB.seq.*
45: /cgn2_6/ptodata/2/pna/US6006.COMB.seq.*
46: /cgn2_6/ptodata/2/pna/US6007.COMB.seq.*
47: /cgn2_6/ptodata/2/pna/US6008.COMB.seq.*
48: /cgn2_6/ptodata/2/pna/US6009.COMB.seq.*
49: /cgn2_6/ptodata/2/pna/US6010.COMB.seq.*
50: /cgn2_6/ptodata/2/pna/US6011.COMB.seq.*
51: /cgn2_6/ptodata/2/pna/US6012.COMB.seq.*
52: /cgn2_6/ptodata/2/pna/US6013.COMB.seq.*
53: /cgn2_6/ptodata/2/pna/US6014.COMB.seq.*
54: /cgn2_6/ptodata/2/pna/US6015.COMB.seq.*
55: /cgn2_6/ptodata/2/pna/US6016.COMB.seq.*
56: /cgn2_6/ptodata/2/pna/US6017.COMB.seq.*
57: /cgn2_6/ptodata/2/pna/US6018.COMB.seq.*
58: /cgn2_6/ptodata/2/pna/US6019.COMB.seq.*
59: /cgn2_6/ptodata/2/pna/US6020.COMB.seq.*
60: /cgn2_6/ptodata/2/pna/US6021.COMB.seq.*
61: /cgn2_6/ptodata/2/pna/US6022.COMB.seq.*
62: /cgn2_6/ptodata/2/pna/US6023.COMB.seq.*
63: /cgn2_6/ptodata/2/pna/US6024.COMB.seq.*
64: /cgn2_6/ptodata/2/pna/US6025.COMB.seq.*
65: /cgn2_6/ptodata/2/pna/US6026.COMB.seq.*
66: /cgn2_6/ptodata/2/pna/US6027.COMB.seq.*
67: /cgn2_6/ptodata/2/pna/US6028.COMB.seq.*
68: /cgn2_6/ptodata/2/pna/US6029.COMB.seq.*
69: /cgn2_6/ptodata/2/pna/US6030.COMB.seq.*
70: /cgn2_6/ptodata/2/pna/US6031.COMB.seq.*
71: /cgn2_6/ptodata/2/pna/US6032.COMB.seq.*
72: /cgn2_6/ptodata/2/pna/US6033.COMB.seq.*
73: /cgn2_6/ptodata/2/pna/US6034.COMB.seq.*
74: /cgn2_6/ptodata/2/pna/US6035.COMB.seq.*
75: /cgn2_6/ptodata/2/pna/US6036.COMB.seq.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1830	100.0	1830	29	US-09-755-456-1
2	176	9.6	1116	29	US-09-755-456-1
3	52.8	2.9	2000	33	US-09-887-272A-5283
4	50.8	2.8	2000	33	US-09-887-272A-5283
5	45.4	2.5	740	1	PCT-US00-14864A-26
6	45.4	2.5	740	35	US-09-950-083-995
7	42.6	2.3	533	57	US-60-181-996-561
8	42.2	2.3	1011	75	US-60-360-207-36034
9	42.2	2.3	1564	35	US-09-950-083-7697
10	41.4	2.3	516	58	US-60-196-718-3736
11	41.4	2.3	516	58	US-60-196-718-3736
12	40.8	2.2	1084	23	US-09-610-154-72
13	40.8	2.2	1084	35	US-09-938-301-72
14	40.6	2.2	7218	8	US-08-466-194-14
15	40	2.2	470	55	US-60-162-842-820
16	40	2.2	470	55	US-60-162-842-820
17	40	2.2	832	24	US-09-621-976-2813
18	40	2.2	832	24	US-09-621-976-2813
19	40	2.2	2791	75	US-60-147-939-2813
20	39.8	2.2	4064	31	US-60-360-207-1380
21	39.8	2.2	4064	31	US-60-360-207-1380
22	39.2	2.1	660	57	US-09-818-656-3
23	39.2	2.1	708	37	US-60-188-162-1150
24	38.8	2.1	1026	37	US-10-001-873-26
25	38.6	2.1	467	23	US-09-605-700-521
26	38.6	2.1	436	19	US-09-528-409-101912
27	38.6	2.1	436	35	US-09-933-524-101912
28	38.6	2.1	436	35	US-09-933-524-101912
29	38.2	2.1	587	29	US-09-721-589-3986
30	38.2	2.1	611	1	PCT-US01-006633-13854
31	38.2	2.1	611	33	US-09-864-761-17426
32	38.2	2.1	611	33	US-09-864-761-17426
33	38.2	2.1	611	33	US-09-864-761-17426
34	38.2	2.1	611	33	US-09-864-761-17426
35	38.2	2.1	611	33	US-09-864-761-17426
36	38.2	2.1	611	33	US-09-864-761-17426
37	38.2	2.1	611	33	US-09-864-761-17426
38	38.2	2.1	611	33	US-09-864-761-17426
39	38.2	2.1	611	33	US-09-864-761-17426
40	38.2	2.1	611	33	US-09-864-761-17426
41	38.2	2.1	611	33	US-09-864-761-17426
42	38.2	2.1	611	33	US-09-864-761-17426
43	38.2	2.1	611	33	US-09-864-761-17426
44	38.2	2.1	611	33	US-09-864-761-17426
45	38.2	2.1	611	33	US-09-864-761-17426
46	38.2	2.1	611	33	US-09-864-761-17426
47	38.2	2.1	611	33	US-09-864-761-17426
48	38.2	2.1	611	33	US-09-864-761-17426
49	38.2	2.1	611	33	US-09-864-761-17426
50	38.2	2.1	611	33	US-09-864-761-17426
51	38.2	2.1	611	33	US-09-864-761-17426
52	38.2	2.1	611	33	US-09-864-761-17426
53	38.2	2.1	611	33	US-09-864-761-17426
54	38.2	2.1	611	33	US-09-864-761-17426
55	38.2	2.1	611	33	US-09-864-761-17426
56	38.2	2.1	611	33	US-09-864-761-17426
57	38.2	2.1	611	33	US-09-864-761-17426
58	38.2	2.1	611	33	US-09-864-761-17426
59	38.2	2.1	611	33	US-09-864-761-17426
60	38.2	2.1	611	33	US-09-864-761-17426
61	38.2	2.1	611	33	US-09-864-761-17426
62	38.2	2.1	611	33	US-09-864-761-17426
63	38.2	2.1	611	33	US-09-864-761-17426
64	38.2	2.1	611	33	US-09-864-761-17426
65	38.2	2.1	611	33	US-09-864-761-17426
66	38.2	2.1	611	33	US-09-864-761-17426
67	38.2	2.1	611	33	US-09-864-761-17426
68	38.2	2.1	611	33	US-09-864-761-17426
69	38.2	2.1	611	33	US-09-864-761-17426
70	38.2	2.1	611	33	US-09-864-761-17426
71	38.2	2.1	611	33	US-09-864-761-17426
72	38.2	2.1	611	33	US-09-864-761-17426
73	38.2	2.1	611	33	US-09-864-761-17426
74	38.2	2.1	611	33	US-09-864-761-17426
75	38.2	2.1	611	33	US-09-864-761-17426

32 38.2 2.1 2019 75 US-60-360-207-17623 Sequence 17623, A
 c 33 38 2.1 1501 31 US-09-815-264-50372 Sequence 50372, A
 c 34 38 2.1 3955 24 US-09-620-392-58240 Sequence 58240, A
 c 35 38 2.1 5044 24 US-09-620-392-65702 Sequence 65702, A
 c 36 38 2.1 8419 24 US-09-620-392-9985 Sequence 9985, Ap
 c 37 38 2.1 10829 24 US-09-620-392-30253 Sequence 30253, A
 c 38 38 2.1 85882 24 US-09-620-392-48150 Sequence 48150, A
 c 39 38 2.1 94641 28 US-09-702-134-7928 Sequence 7928, Ap
 c 40 38 2.1 94641 31 US-09-815-264-79401 Sequence 79401, A
 c 41 37.8 2.1 678 24 US-09-634-306B-121331 Sequence 121331, A
 c 42 37.8 2.1 771 24 US-09-634-306B-150013 Sequence 150013, A
 c 43 37.8 2.1 2727 33 US-09-887-272A-3090 Sequence 3090, Ap
 c 44 37.8 2.1 2727 71 US-60-325-448-694 Sequence 694, App
 c 45 37.6 2.1 411 16 US-09-293-972-31298 Sequence 31298, A

ALIGNMENTS

RESULT 1
 ; Sequence 1, Application US/09755456
 ; GENERAL INFORMATION:
 ; APPLICANT: DELBAC, FREDERIC
 ; APPLICANT: DANCHIN, ANTOINE
 ; APPLICANT: VIVARES, CHRISTIAN
 ; TITLE OF INVENTION: MICROSPORIDIAN POLAR TUBE PROTEINS, NUCLEIC ACIDS
 ; TITLE OF INVENTION: CODING FOR THESE PROTEINS AND THEIR APPLICATIONS
 ; FILE REFERENCE: 1566-00
 ; CURRENT APPLICATION NUMBER: US/09/755,456
 ; CURRENT FILING DATE: 2001-01-05
 ; PRIOR APPLICATION NUMBER: PCT/FR99/01630
 ; PRIOR FILING DATE: 1999-07-06
 ; PRIOR APPLICATION NUMBER: FR 98/08692
 ; PRIOR FILING DATE: 1998-07-07
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 1830
 ; TYPE: DNA
 ; ORGANISM: Encephalitozoon cuniculi
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (345)..(1529)
 ; US-09-755-456-1

Query Match 100.0%; Score 1830; DB 29; Length 1830;
 Best Local Similarity 100.0%; Pred. NO. 0;
 Matches 1830; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gaattcagatgctcatctacattggaattaaatgatgatcttattgtatataatcctg 60
 Db 1 gaattcagatgctcatctacattggaattaaatgatgatcttattgtatataatcctg 60
 QY 61 ggcgcagcagcgccgctgctgattcttcaggggtgcgcctaccagtgacagagggtcc 120
 Db 61 ggcgcagcagcgccgctgctgattcttcaggggtgcgcctaccagtgacagagggtcc 120
 QY 121 ggaaggtcttgatgaaagtaaggccattgtgtggtcttcacatgcatcagtcgcct 180
 Db 121 ggaaggtcttgatgaaagtaaggccattgtgtggtcttcacatgcatcagtcgcct 180
 QY 181 ttccggtcttccacaagatcccaattatctccccaagacttccaacctcagaatgaaa 240
 Db 181 ttccggtcttccacaagatcccaattatctccccaagacttccaacctcagaatgaaa 240
 QY 241 cagagatgaactctctgtgcaaatctgataatcgaattggagacattgaaccacgag 300
 Db 241 cagagatgaactctctgtgcaaatctgataatcgaattggagacattgaaccacgag 300
 QY 301 ttggaataaagtataatactccgaaaacgcagagtttaagatgaagaagtattctta 360
 Db 301 ttggaataaagtataatactccgaaaacgcagagtttaagatgaagaagtattctta 360

Db 301 ttggaataaagtataatactccgaaaacgcagagtttaagatgaagaagtattctta 360
 QY 361 agatctctctgctctattgcccctgatgaagtggaaagtgtcttaagcaaccgac 420
 Db 361 agatctctctgctctattgcccctgatgaagtggaaagtgtcttaagcaaccgac 420
 QY 421 tgtgcagcaatgataatgacctcctcgggaaacaaaggtatgtgtcagcgcgtcgt 480
 Db 421 tgtgcagcaatgataatgacctcctcgggaaacaaaggtatgtgtcagcgcgtcgt 480
 QY 481 atgtgtgtatcccaagaccaccgggaaacaaatagcaaacgtgtcaagcgtttcaagaa 540
 Db 481 atgtgtgtatcccaagaccaccgggaaacaaatagcaaacgtgtcaagcgtttcaagaa 540
 QY 541 catattctctctcctccgctgaccccaatctcccaatgtgctccgggaaactaagaga 600
 Db 541 catattctctctcctccgctgaccccaatctcccaatgtgctccgggaaactaagaga 600
 QY 601 atgagacatctccatcgtgctctcgcagaaagtgaagaaatgcaagattgcgtatga 660
 Db 601 atgagacatctccatcgtgctctcgcagaaagtgaagaaatgcaagattgcgtatga 660
 QY 661 agcactgcgacgcagcagcaagaacaacatacaggaagacacccaggtacagccttgtaa 720
 Db 661 agcactgcgacgcagcagcaagaacaacatacaggaagacacccaggtacagccttgtaa 720
 QY 721 ccccaagcagcaagcctctgtcagtgatctcccaactctcgcgtacccgtgactg 780
 Db 721 ccccaagcagcaagcctctgtcagtgatctcccaactctcgcgtacccgtgactg 780
 QY 781 tggagtctgcaagcttcacatctgtgtgcaggtgttcctgtcgtgtcaccacagag 840
 Db 781 tggagtctgcaagcttcacatctgtgtgcaggtgttcctgtcgtgtcaccacagag 840
 QY 841 cagttccagcctactatacaacaatggaactccggtatcttcctgagacagcaagatcttt 900
 Db 841 cagttccagcctactatacaacaatggaactccggtatcttcctgagacagcaagatcttt 900
 QY 901 ctgacactctcccccagcagcactctgtgtcagggagacagccatgcttaacactctg 960
 Db 901 ctgacactctcccccagcagcactctgtgtcagggagacagccatgcttaacactctg 960
 QY 961 gacagcaacagatcctttctggaactctcccccaggggtcaactgtgtcagggagcag 1020
 Db 961 gacagcaacagatcctttctggaactctcccccaggggtcaactgtgtcagggagcag 1020
 QY 1021 ccaagcctagcactcctctgagcagcaagtccttctgtgacactctcccccagagatca 1080
 Db 1021 ccaagcctagcactcctctgagcagcaagtccttctgtgacactctcccccagagatca 1080
 QY 1081 ctttgtcagggagcagcagcagccttagcactctggtgagcaacagcgtcttcttgaga 1140
 Db 1081 ctttgtcagggagcagcagcagccttagcactctggtgagcaacagcgtcttcttgaga 1140
 QY 1141 ccccttcccaaggagcactctgtgtcaggaatcaaggtatgctctggaacatcccgagttc 1200
 Db 1141 ccccttcccaaggagcactctgtgtcaggaatcaaggtatgctctggaacatcccgagttc 1200
 QY 1201 ctggaacagcagggagcagcttagtgacagtggtgtgcccctcaagattccaaccctgtca 1260
 Db 1201 ctggaacagcagggagcagcttagtgacagtggtgtgcccctcaagattccaaccctgtca 1260
 QY 1261 tgcgcgcatccatgaacatagtgaaagggatccctctctcaacgatacaagccaa 1320
 Db 1261 tgcgcgcatccatgaacatagtgaaagggatccctctctcaacgatacaagccaa 1320
 QY 1321 accctgataactggatcctctgtgtgacatacagaagcgggggagacatcctctgagc 1380
 Db 1321 accctgataactggatcctctgtgtgacatacagaagcgggggagacatcctctgagc 1380
 QY 1381 aaaaaaccgagaagttccgcacagcagtatgacatggaagcctgtgcaaacacaaacaa 1440
 Db 1381 aaaaaaccgagaagttccgcacagcagtatgacatggaagcctgtgcaaacacaaacaa 1440

QY	1441	cggtattattataggcaagaagagatctctgtttggaccaggaatgtcaaatgtcaat	1500
Db	1441	cggtattattataggcaagaagagatctctgtttggaccaggaatgtcaaatgtcaat	1500
QY	1501	ctccatgtcaaacactgcgtctccaaatgtctgtccaaagctaaataaagaagtttaattctt	1560
Db	1501	ctccatgtcaaacactgcgtctccaaatgtctgtccaaagctaaataaagaagtttaattctt	1560
QY	1561	tctcttcggtctcttttggaaagcttggatltgggagatctaaatgtgggctgaagtgaaat	1620
Db	1561	tctcttcggtctcttttggaaagcttggatltgggagatctaaatgtgggctgaagtgaaat	1620
QY	1621	gccaacacctctctctgcccaagaacacatctggagtgctctctcttggccagagtttg	1680
Db	1621	gccaacacctctctctgcccaagaacacatctggagtgctctctcttggccagagtttg	1680
QY	1681	taacagaagattcccccagagagatttagacagcctttgagatcacatgttaagatcagattaaac	1740
Db	1681	taacagaagattcccccagagagatttagacagcctttgagatcacatgttaagatcagattaaac	1740
QY	1741	tctctcaaatattattatctctctctgtttatataaccagagcaatcttgaagaatgtgctc	1800
Db	1741	tctctcaaatattattatctctctgtttatataaccagagcaatcttgaagaatgtgctc	1800
QY	1801	gaattcaagcttcctctagaagctgtggatc	1830
Db	1801	gaattcaagcttcctctagaagctgtggatc	1830

```

RESULT 2
US-09-755-456-3
; Sequence 3, Application US/09755456
; GENERAL INFORMATION:
; APPLICANT: DELBAC, FREDERIC
; APPLICANT: DANCHIN, ANTOINE
; APPLICANT: VIVARES, CHRISTIAN
; TITLE OF INVENTION: MICROSPORDIAN POLAR TUBE PROTEINS, NUCLEIC ACIDS
; TITLE OF INVENTION: CODING FOR THESE PROTEINS AND THEIR APPLICATIONS
; FILE REFERENCE: 1566-00
; CURRENT APPLICATION NUMBER: US/09/755,456
; CURRENT FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: PCT/FR99/01630
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: FR 98/08692
; PRIOR FILING DATE: 1998-07-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1116
; TYPE: DNA
; ORGANISM: Eucephalitozoon intestinalis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1113)
; US-09-755-456-3

```

Query Match	9.68;	Score 176;	DB 29;	Length 1116;
Best Local Similarity	53.58;	Pred. No. 6.6e-42;		
Matches 635;	Conservative 0;	Mismatches 460;	Indels 93;	Gaps 9

Accession	Sequence	Position
Oy	atgaaggatcttctctaaagacccctctgcgcctctattgcgcctgataagaaatttggagaaatgc	404
345		
Oy	atgaaggatcttctctaaagacccctctgcgcctctattgcgcctgataagaaatttggagaaatgc	404
345		
Db	atgaaggatcttctctaaagacccctctgcgcctctattgcgcctgataagaaatttggagaaatgc	60
1		
Oy	tatttcagaacacgcacgctgtgcgaagcaatgatgacctaatcccgagacaacaagagtatg	466
405		
Db	tatttcagaacacgctgtgcgaagcaatgatgacctaatcccgagacaacaagagtatg	109
61		
Oy	tatttcagaacacgctgtgcgaagcaatgatgacctaatcccgagacaacaagagtatg	524
465		
Db	tatttcagaacacgctgtgcgaagcaatgatgacctaatcccgagacaacaagagtatg	168
110		

[illegible]

RESULT 3
US-09-887-272A-5263
; Sequence 5263, Application US/09887272A

```

: GENERAL INFORMATION:
: APPLICANT: Hou, Yu-Ming
: APPLICANT: Quan, Sheng
: APPLICANT: Chang, Hur-Song
: APPLICANT: Zhu, Tong
: APPLICANT: Whitham, Steve
: APPLICANT: Goff, Steve
: APPLICANT: Glazebrook, Jane
: APPLICANT: Chen, Wengulong
: APPLICANT: Katagiri, Fumaki
: APPLICANT: Xie, Zhiyi
: APPLICANT: Tao, Yi
: APPLICANT: Zou, Guangzhou
: APPLICANT: Cooper, Bret
: TITLE OF INVENTION: PLANT GENES INVOLVED IN DEFENSE AGAINST
: TITLE OF INVENTION: PATHOGENS
: FILE REFERENCE: 1360.003052
: CURRENT APPLICATION NUMBER: US/09/887,272A
: CURRENT FILING DATE: 2001-06-23
: PRIOR APPLICATION NUMBER: 60/213,634
: PRIOR FILING DATE: 2000-06-23
: PRIOR APPLICATION NUMBER: 60/214,926
: PRIOR FILING DATE: 2000-06-23
: PRIOR APPLICATION NUMBER: 60/261,320
: PRIOR FILING DATE: 2001-01-12
: PRIOR APPLICATION NUMBER: 60/264,353
: PRIOR FILING DATE: 2001-01-26
: PRIOR APPLICATION NUMBER: 60/273,879
: PRIOR FILING DATE: 2001-03-07
: PRIOR APPLICATION NUMBER: 09/887,271
: PRIOR FILING DATE: 2001-06-22
: NUMBER OF SEQ ID NOS: 6813
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 5263
: LENGTH: 2000
: TYPE: DNA
: ORGANISM: Oryza sativa
: US-09-887-272A-5263

```

Query Match	Similarity	2.9%	Score 52.8:	DB 33:	Length 2000:
Best Local	Similarity	9.1%	Pred. No. 0.00046:		
Matches	93:	Conservative	457:	Mismatches	464:
				Indels	6:
				Gaps	3
QY	773	ggtgacgtgtgagcttcgcacagcttcacatctgtgtgcacagttgttcgttcgtca	832		
Db	7	gwtscswrrgyrrimymagmascarmsarmrkngsmaskykccsacgckmtrrkxkw	66		
QY	833	ccacacagcagttcagcagctactaacacatggaacatccggtatctccgcgaacaa	892		
Db	67	ysaasasrptgtgkwsasgsygtgmkkrktrkrwrgrrgmrrsmrmgyrricarsg	126		
QY	893	gattctttcgtgcaactctcccccagggacgactttgttcagggacagcatgctag	952		
Db	127	rmaagsgymmgkksrmsywwmcyargcgckrkksksgwgttcrrgargsgwsagaky	186		
QY	953	cactctctggacagcaacagatcctcttcctggcactcttccccaggggtcaacttgytca	1012		
Db	187	ksgsmksrkmmmsccgrsgcgrrrasryrygtarkygltkmlysasrsmraymtsyaw	246		
QY	1013	ggagacagcgaagcctagcactcctcggcgagcaacagagtccttctcggcacttccccc	1072		
Db	247	acssytwcrskrrismmwkmrmkrmvrsrsgywsykhmmctcaykksysrwy----	302		
QY	1073	aggagtcactttgtgtcagaagacagagccagcctcctagcactctcgtggcagaacagctct	1132		
Db	303	rggvrgalrlywrgyismrmannmykkmlyrygylgmkrgmwagrmmmrmcwkscakym	362		
QY	1133	tctcggacccctcccccagggacgcaacttgtgtcagagtaagatcctctgacatc	1192		
Db	363	rwmrmrmlrrrrwakkssrtstrkrkrkwcmrkrykymgyrsmisckrtaywmktrcgsr	422		
QY	1193	cggagttcctcgacagcaagacagcttagtbgaaacgtgtgtgcctccatgattccaa	1252		

Db	423	awkingcgcgcmctcmksygmrmwksvkrmskkykwmrmymrwlkckcstttmgtkrrgmm	482
Qy	1253	ccctgtcatgcgcgcacatccatagacatagtagtgaatggatctcttc-accgcat	1311
Db	483	gtmgrrcykkrrsgmkkrerrrrwgrmyrmwkrmysarylmrlyarkkysyaarkc	542
Qy	1312	acagcccaaacctcgcgatctcgtggtctctgtgttgacataagaagacggggggaacat	1371
Db	543	wyrjkgtywagwmwmykrykrmymyknmmwmykrykskcswwkmsyyscmksarkkagkmc	602
Qy	1372	ccttcgagcgaacaaacccgagagatccgcacgacgatcgaatgacatggagccttgcaaac	1431
Db	603	krskmsawksmsrstrkrcxkaskkrskakrkyamngmbsgsmsrkwstycywkrgsmk	662
Qy	1432	caacacaaacgggtatattatagagaacagcgagatctcttgacacgaagtatacatg	1491
Db	663	stclmwmymmkkyakkygswyryrlyawcmymrwyryysymlymawylsstmatngt	722
Qy	1492	caattaactctccatgaacacatgctgtccaatgctgactagctaaataaaacgagtt	1551
Db	723	mkysgyrwtswykyckcsww-kyrsmmyywswwmaklwmrlyatrrmmwmyrysmkwt	781
Qy	1552	aactctctttctctgctctcttttggaacgttgagggagggagggagctatgsgctg	1611
Db	782	wclmwmymwmywtlytmrmymwmykcktklywysaatlygtlwaaawmaktkmrgmtgkt	841
Qy	1612	aaggaatgcacacactctctccccaagaacacatcgagatgtctctcctgsgca	1671
Db	842	rgrrkarerymwkwatwatcktrmmltkgkgaawltmakaewrlyywswnawyyktrrt	901
Qy	1672	ggaagtttggaacagagattcccgagagattlagcagccttgaglaacatgattgaatca	1731
Db	902	rytkcwwkarkarwsgwayrmmwkwsgakmmwmmkgyrgwvckkywyyctlkmacgatkym	961
Qy	1732	gtataactctccaattattttattcttcttgttatcccgagcgaatccgaa	1791
Db	962	ccagwamysyswrltymtlwmtlwmaasrtakrmarmmktrawsksyataywkmagca	1021
RESULT 4			
US-09-887-272A-5263/C			
Sequence 5263, Application US/09887272A			
GENERAL INFORMATION:			
APPLICANT: Hou, Yu-Ming			
APPLICANT: Quan, Sheng			
APPLICANT: Chang, Hur-Song			
APPLICANT: Zhu, Tong			
APPLICANT: Whitlam, Steve			
APPLICANT: Golf, Steve			
APPLICANT: Glazebrook, Jane			
APPLICANT: Chen, Wengqiong			
APPLICANT: Katagiri, Fumiaki			
APPLICANT: Xie, Zhiyi			
APPLICANT: Tao, Yi			
APPLICANT: Zou, Guangzhou			
APPLICANT: Cooper, Bret			
TITLE OF INVENTION: PLANT GENES INVOLVED IN DEFENSE AGAINST			
TITLE OR INVENTION: PATHOGENS			
FILE REFERENCE: 1360.003US2			
CURRENT APPLICATION NUMBER: US/09/887, 272A			
CURRENT FILING DATE: 2001-06-23			
PRIOR APPLICATION NUMBER: 60/213, 634			
PRIOR FILING DATE: 2000-06-23			
PRIOR APPLICATION NUMBER: 60/214, 926			
PRIOR FILING DATE: 2000-06-23			
PRIOR APPLICATION NUMBER: 60/261, 320			
PRIOR FILING DATE: 2001-01-12			
PRIOR APPLICATION NUMBER: 60/264, 353			
PRIOR FILING DATE: 2001-01-26			
PRIOR APPLICATION NUMBER: 60/273, 879			
PRIOR FILING DATE: 2001-03-07			
PRIOR APPLICATION NUMBER: 09/887, 271			

```

1 RESULT 4
2 US-09-887-272A-5263/c
3 : Sequence 5263, Application US/09887272A
4 : GENERAL INFORMATION:
5 : APPLICANT: Hou, Yu-Ming
6 : APPLICANT: Quan, Sheng
7 : APPLICANT: Chang, Hur-Song
8 : APPLICANT: Zhu, Tong
9 : APPLICANT: Whitlam, Steve
10 : APPLICANT: Golf, Steve
11 : APPLICANT: Glazebrook, Jane
12 : APPLICANT: Chen, Wengulong
13 : APPLICANT: Katagiri, Fumitsuki
14 : APPLICANT: Xie, Zhiyi
15 : APPLICANT: Tao, Yi
16 : APPLICANT: Zou, Guangzhou
17 : APPLICANT: Cooper, Bret
18 : TITLE OR INVENTION: PLANT GENES INVOLVED IN DEFENSE AGAINST
19 : TITLE OF INVENTION: PATHOGENS
20 : FILE REFERENCE: 1360.003US2
21 : CURRENT APPLICATION NUMBER: US/09/887, 272A
22 : CURRENT FILING DATE: 2001-06-23
23 : PRIOR APPLICATION NUMBER: 60/213, 634
24 : PRIOR FILING DATE: 2000-06-23
25 : PRIOR APPLICATION NUMBER: 60/214, 926
26 : PRIOR FILING DATE: 2000-06-23
27 : PRIOR APPLICATION NUMBER: 60/261, 320
28 : PRIOR FILING DATE: 2001-01-12
29 : PRIOR APPLICATION NUMBER: 60/264, 353
30 : PRIOR FILING DATE: 2001-01-26
31 : PRIOR APPLICATION NUMBER: 60/273, 879
32 : PRIOR FILING DATE: 2001-03-07
33 : PRIOR APPLICATION NUMBER: 09/887, 271

```

```

:
: PRIOR FILING DATE: 2001-06-22
: NUMBER OF SEQ ID NOS: 6813
: SOFTWARE: FASTSEQ for Windows Version 4.0.
: SEQ ID NO 5263
: LENGTH: 2000
: TYPE: DNA
: ORGANISM: Oryza sativa
: US-09-887-272A-5263

```

Query Match	2.8%	Score 50.8	DB 33	Length 2000
Best Local Similarity	9.4%	Pred. No. 0.0019		
Matches	89	Conservative 424	Mismatches 426	Indels 7
				Gaps 3

QY	207	atctctccaagacttccaaccctcagatcggaaacagatgaatcaactctctgtgcaac	266
Db	958	ATYCGTMMMAAGNRHMAACWYCCMMKKMMKMTSCMMKKYMTNWSCTTMMGAAIRYAY	899
QY	267	gtagatctgacttggagacattgaaccacgaagtttgaataaagataataactctc	326
Db	988	AMRRRRRTYTWMSRRMYTWTMTKMAMTMMCMCAKMYNATGTATMMMYTYTYCYAMT	839
QY	327	gaaacgcagag-cttaagcttgaagtaattcttaagaccctctgtcctcatggcct	385
Db	838	CACKKYTMAMTMMTMMTWACMRAITSRRWRALMAGMRKKRTKMKRAYMTWRMRCKMKGARW	779
QY	386	gatgaacttggaaatgtctattcagaacaacgcactgttgaacaaatgatagtactac	445
Db	778	MKSRYPKKMKKRYATRYMKMMAMTMMMSRRMKSYPHMGMRMRMSAIVRSRMAKCAKT	719
QY	446	tcggcgcaacaaaggtatgctcagaagccgctgtatgtctgatalccccaagaccggg	505
Db	718	KYASASAMTKRARRSYRRRRMYMKKKGMYYRYSRCSRRMRARRSKRRKMAAGSMSC	659
QY	506	aaccatgacaacttgtcgaacgcggttcacagaacacatactctccttcccgcgcac	565
Db	658	WMYRGARASMMYSKYSCSAAKCKKTYTMTSYMS-----TGYTGMYSYKSMSTSKMSY	604
QY	566	cacactccagtgactccggggaaacatcagcgaatgagacatctccatcgctcctgc	625
Db	603	MGMKTCMTMYSMKSTFRSMKGMSSOMSRMYMMKMKRRMYRMKMKCTWRRCMYR	544
QY	626	agaagatgttagaacatgacaatltgcglattgaagacgttcgaagcccaagaaac	685
Db	543	MGTYMTTSSRRMYTGRTYARTSKRRRTMYKRYKCYTYTGYMTKCSIMRGTCKA	484
QY	686	atcagggagacacacagagtgacggccttgttgaaccacaagcagcagaacagcttgc	745
Db	483	CKKCCYAMCKAAVSGMMMYHRYKSKMMBMSTKYWMSMWYKCKSMKYGAKGCGCKMW	424
QY	746	agtatctccaacactctcgcgttaccggtgactgtgagctgtcgaacttcacagctc	805
Db	423	TYCGYGMKMYTYMGSYKYSRCYKRYRMMYTKGMMTYMTYSAVSSMMTYTYTYAYTYMWY	364
QY	806	tgltccaagtgttccctctgtctgctccaacacacagcagtlcagaactatacaacatg	865
Db	363	YKRGTMSWYGSYKXKXCYTMMCYMCMCHYRRKMKMRKKTTYSKRCYMRATATYMWCC	304
QY	866	aacatccggtatctccttggagacaacaagatccttcttgcgaactctccccaagag	924
Db	303	YRKGMYSRSMAMNTACKMMRSMWSMBCRSYSYMYKMKMMKKSYYMSGARMSGRTWSR	244
QY	925	cttggltcagggacagcgcatgctcagaacctcttgaacagaacagatccttcttgca	984
Db	243	SAAKRTYKGYSTRAKMMAACRMYSACRISTNTSYCCGCTCGSSKMYMKSCSMRRT	184
QY	985	ctcttccccaagggatcacttltgttcaggaagcagcagcgtacgaactcttggcagc	1044
Db	183	CSMSCCSYCYCYAMCMSCMSMYSMWGSCGYTGBMKMRSKYSMCKKYGSCGCRKYSY	124
QY	1045	aacaggtccttctcgggaactcttccccaagagatcaacttgtctcaggaacagccaagc	1104
Db	123	TGYRKYCKMKWYSYKCYCYCYWYMSYMYMKCMCSHSCSSWMSCAVCTSSYSSRWMS	64

Oy 1105 ctgacactcctggcgagcaaacagtcccttctgcacccttcgcc 1150
::|::|::|::|:
Db 63 MYAAKMGMCSSGMYRMSKSCKMYSKYSSCAKTGSKKCTKRKYY 18

```

RESULT      5
PCT-US00-14964A-26
: Sequence 26, Application PC/7US0014964A
: GENERAL INFORMATION:
: APPLICANT: Human Genome Sciences, Inc.
: TITLE OF INVENTION: 48 Human Secreted Proteins
: FILE REFERENCE: P5560PCT
: CURRENT APPLICATION NUMBER: PCT/US00/14964A
: CURRENT FILING DATE: 2000-06-01
: PRIOR APPLICATION NUMBER: 60/138,627
: PRIOR FILING DATE: 1999-06-11
: NUMBER OF SEQ ID NOS: 170
: SOFTWARE: Patent Ver. 2.0
: SEQ ID NO 26
: LENGTH: 740
: type: DNA
: ORGANISM: Homo sapiens
PCT-US00-14964A-26

```

Query Match	2.58	Score 45.4	DB 1	Length 740
Best Local Similarity	47.48	Pred. No.	0.048	
Matches 136, Conservative	0	Mismatches 151,	Indels 0	Gaps 0

[illegible]

```

1 RESULT 6
2 US-09-950-083-995
3 ; Sequence 995, Application US/09950083
4 ; GENERAL INFORMATION:
5 ; APPLICANT: Rosen, et. al
6 ; TITLE OF INVENTION: Human Secreted Proteins
7 ; FILE REFERENCE: PS805
8 ; CURRENT APPLICATION NUMBER: US/09/950, 083
9 ; CURRENT FILING DATE: 2001-09-12
10 ; PRIOR APPLICATION NUMBER: 60/278, 650
11 ; PRIOR FILING DATE: 2001-03-27
12 ; PRIOR APPLICATION NUMBER: 09/833, 245
13 ; PRIOR FILING DATE: 2001-04-12
14 ; PRIOR APPLICATION NUMBER: PCT/US01/11988
15 ; PRIOR FILING DATE: 2001-04-12
16 ; PRIOR APPLICATION NUMBER: PCT/US00/06043
17 ; PRIOR FILING DATE: 2000-03-09
18 ; PRIOR APPLICATION NUMBER: PCT/US00/06012
19 ; PRIOR FILING DATE: 2000-03-09
20 ; PRIOR APPLICATION NUMBER: PCT/US00/06058
21 ; PRIOR FILING DATE: 2000-03-09
22 ; PRIOR APPLICATION NUMBER: PCT/US00/06044

```


QY 1122 caacag 1128
DB 172 GCACAG 166

RESULT 9

US-09-950-083-7697
Sequence 7697, Application US/09950083
GENERAL INFORMATION:
APPLICANT: Rosen, et. al
TITLE OF INVENTION: Human Secreted Proteins
FILE REFERENCE: PS805
CURRENT APPLICATION NUMBER: US/09/950,083
PRIOR FILING DATE: 2001-09-12
PRIOR APPLICATION NUMBER: 60/278,650
PRIOR FILING DATE: 2001-03-27
PRIOR APPLICATION NUMBER: 09/833,245
PRIOR FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: PCT/US01/11988
PRIOR FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: PCT/US00/06043
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06012
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06058
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06044
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06059
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06042
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06014
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06013
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06049
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06057
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06824
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06765
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06792
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06830
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06782
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06822
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06791
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06828
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06823
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06781
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/07505
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07440
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07506
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07507
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07535
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07525
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07534

PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07483
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07526
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07527
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07661
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07579
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07723
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07724
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/14929
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/07722
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07578
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07726
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07677
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07725
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/09070
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/08982
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/08983
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/09067
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/09066
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/09068
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/08981
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/08980
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/09071
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/09069
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/15136
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/14926
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/14963
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/15135
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/14934
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/14933
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/15137
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/14928
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/14973
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/14964
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/26376
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: PCT/US00/26371
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: PCT/US00/26324


```
; CURRENT FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 7494
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3736
; LENGTH: 516
; TYPE: DNA
; ORGANISM: HUMAN
US-60-196-718-3736
```

```
Query Match          2.3%; Score 41.4; DB 58; Length 516;
Best Local Similarity 45.7%; Pred. No. 0.66;
Matches 144; Conservative 0; Mismatches 171; Indels 0; Gaps 0;
```

```
QY 863 tggacatccggtattcttgagacaagatccttctgagcacttcccccagagc 922
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 121 tggcgtccctcactcgtgtacacatgagcagcgctggtcgtctcactctgt 180
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 923 cacttgtgtcaggaagcagcatgctcctagcactctgagacaagaatccttctg 982
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 181 gcatatgagagcagagccttgcgtgcctcactgctgtcacatgtgagtcgagcctg 240
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 983 cacttccccccaggggtcacttctgtgtcaggaagcagcgctcctcctggca 1042
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 241 cactgcccctcactgctgtcagcgtgagcctgagcctgcgcctcactgtgtgca 300
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1043 gcaacaggtccttcttgacacttcccccagagtcacttctgtcaggaagcagccac 1102
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 301 cgtgagcctgagcctgcgtgcctcctcactgctgtcagcgtgtgagctgtgagcctg 360
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1103 gcttagcactcctcgtgagacaaggtccttcttgacaccccttcccaaggagcactt 1162
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 361 tgcctcactcgtcgtcagcgtgagcctgagcctgctgcctcactgtgtgacgt 420
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1163 gtgtcaggaatcaag 1177
    | | | | | | | |
DB 421 ggaagctgagcctgtg 435
```

```
RESULT 12
US-09-610-154-72
; Sequence 72, Application US/09610154
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Novel Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-155
; CURRENT APPLICATION NUMBER: US/09/610,154
; CURRENT FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 1084
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-610-154-72
```

```
Query Match          2.2%; Score 40.8; DB 23; Length 1084;
Best Local Similarity 49.5%; Pred. No. 1.5;
Matches 105; Conservative 0; Mismatches 107; Indels 0; Gaps 0;
```

```
QY 905 cacttccccccaggaacacttctgtcaggaagcagcactcctagcactcttgaca 964
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 562 cagcttcagcgccgccaatgtggtctggtcagggcagctcccaactatgcca 621
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 965 gcaacagatccttcttgacacttcccccaggggtcacttctgtcaggaagcagccac 1024
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 622 cagcgtctgtgctcgcagagcttcccaagaaatgacacacatggtctgtcgaagccag 681
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1025 gcttagcactcctcgtgagcaggaaggtccttcttgacacttcccccaggaagcactt 1084
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 682 tcttctggttttagtaacaacaagcttagctcagggcagctcctctgtttacagtcagcat 741
```

```
QY 1085 gtgtcaggaagcagccagcctagcactcctg 1116
    | | | | | | | | | | | | | | | | | |
DB 742 ggaatcgtcagcagctcactcctctggtcagcg 773
```

```
RESULT 13
US-09-938-301-72
; Sequence 72, Application US/09938301
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Novel Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-155
; CURRENT APPLICATION NUMBER: US/09/938,301
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: 09/610,154
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 1084
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-938-301-72
```

```
Query Match          2.2%; Score 40.8; DB 35; Length 1084;
Best Local Similarity 49.5%; Pred. No. 1.5;
Matches 105; Conservative 0; Mismatches 107; Indels 0; Gaps 0;
```

```
QY 905 cacttccccccaggaacacttctgtcaggaagcagcactcctagcactccttgaca 964
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 562 cagcttcagcgccgccaatgtggtctggtcagggcagctcccaactatgcca 621
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 965 gcaacagatccttcttgacacttcccccaggggtcacttctgtcaggaagcagccac 1024
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 622 cagcgtctgtgctcgcagagcttcccaagaaatgacacacatggtctgtcgaagccag 681
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1025 gcttagcactcctcgtgagcaggaaggtccttcttgacacttcccccaggaagcactt 1084
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 682 tcttctggttttagtaacaacaagcttagctcagggcagctcctctgtttacagtcagcat 741
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1085 gtgtcaggaagcagccagcctagcactcctg 1116
    | | | | | | | | | | | | | | | | | |
DB 742 ggaatcgtcagcagctcactcctctggtcagcg 773
```

```
RESULT 14
US-08-466-194-14
; Sequence 14, Application US/08466194
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, NW, Suite 500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20007
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,194
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/232,463
; FILING DATE: 22-APR-1994
```



```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/935,313
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/201 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PTZgpt-F1s
; US-08-466-194-14

```

Query Match 2.2%; Score 40.6; DB 8; Length 7218;

Best Local Similarity 4.7%; Pred. No. 5.1; Matches 19; Conservative 21; Mismatches 175; Indels 0; Gaps 0;

```

QY 746 agtgatctcacacactctccgtaacggtgagctgacagtcacagtcacatctgt 805
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1026 ATTAAATCCAGCTTGCTCGAGCTCGAGGAGCTCGATYYYYYYYYYYYYYY 1085
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 806 tgtaccagttgtctctgtctgtctacacacagcagtcagtcagtcacacatg 865
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1086 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1145
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 866 aacatccgatactctgacagacagatctctctgacatctcccccagagccac 925
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1146 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1205
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 926 ttgtgtcagaggaagcagcagtcagtcagtcagtcagtcagtcagtcagtc 985
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1206 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1265
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 986 tctcccccaggggtcactgtgtcagggacagcagcagtcagtcagtcagtcag 1045
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1266 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1325
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1046 acaggtctctctgacatctctcccccagagagtcacttgttcagggacagcc 1105
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1326 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1385
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1106 tagcacctcgggacagcagcagtcctctctgacacccctcacc 1150
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1386 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1430
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 15

US-60-162-247-820

; Sequence 820, Application US/60162247

; GENERAL INFORMATION:

; APPLICANT: Bonazzi, Vivien

; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND

; TITLE OF INVENTION: USES THEREOF

; FILE REFERENCE: C1000127

; CURRENT APPLICATION NUMBER: US/60/162,247

; NUMBER OF SEQ ID NOS: 5442

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 820

; LENGTH: 470

```

; TYPE: DNA
; ORGANISM: Human
; US-60-162-247-820

```

Query Match 2.2%; Score 40; DB 55; Length 470;

Best Local Similarity 48.7%; Pred. No. 1.7;

Matches 109; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

```

QY 890 acagatctctctgacatctctcccccagagcagtccttggtacggagccatg 949
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 197 acactccctggtctcactctccctggtcacactctccctggtgacatctcc 256
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 950 tagcactctgacagcagacagatctctctgacatctcccccaggggtcact 1009
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 257 tgcacatccctgggtacacactctccctgggtccatctccctggtccgctcc 316
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1010 tcaggagacagccagcctagcactctggtgacagacaggtctctctgac 1069
   || || || || || || || || || || || || || || || || || || || ||
DB 317 attcacatctccctggtgacactccctgggtcacactctccctggtccact 376
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1070 cccaggtcactctgtgtcagggacagccagcctagcactc 1113
   || || || || || || || || || || || || || || || || || || || ||
DB 377 tggggacactctccctggtctcactctccctggtccactc 420
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: September 1, 2002, 03:07:10
Job time: 8529 sec


```
Db 385 SSSSSSSSSYYTSTNC.T.CC...T.MCAABCSYTTTTTTTTT.HSCC.SA..A. 326
OY 1099 ccaagcctagcagctcctgggagcaacagagtccttctggcagcccttcccaaggagca 1158
Db 325 M.YC.A.SYSYS.S.S.S.SYMR.HRA.SHYTRS..S.MCY.YM.Y.Y.Y.YSY 266
OY 1159 cttgtgctagcagatcagatgcttgcgaacatccggagtccttgcgaagcagagcagt 1218
Db 265 YCSRKTM...TMTDM.T.T..MHMY.KYB.HCHTKCRAAT.MN.HTB...N..HBHB.B 206
OY 1219 ctatgtagcagtgctgtgccccctcagattccaacccctgtcatgcccagcatcagaca 1278
Db 205 H.H..B.H.HSNS...TTS...M.TTM.B.TBASH.RARC.STMM.T.MMM.H.S.Y. 146
OY 1279 ttagtgaatgggtatccttctctccgcatacagcccaacccctgagcactggagt 1338
Db 145 .CCR..H.R..R.G...S.CT.HN.B.CYR.RNGMY.HS...S.AA...AHH..DH. 86
OY 1339 cctgtgtgacatacagaagcgg 1362
Db 85 DTYBH.YH.KNNAHC.W.KKNB.S 62
```

```
RESULT 5
US-10-137-871-10/c
; Sequence 10, Application US/10137871
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Audrey E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C153
; CURRENT APPLICATION NUMBER: US/10/137,871
; PRIOR APPLICATION REMOVED - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-137-871-10
```

```
Query Match 2.28; Score 40.6; DB 7; Length 594;
Best Local Similarity 6.98; Pred. No. 0.14;
Matches 35; Conservative 183; Mismatches 286; Indels 0; Gaps 0;
```

```
OY 859 acaatggaacatccggtatctctgacagcaacagatccttctggcacttcccccag 918
Db 565 A.MCTT..AMM.M.CBT.STT.YA.M.YT.S.S.S.SYSYS.S.S.S.SDSYSYA.SY 506
OY 919 gaggcacttctgtcaggaagcagcatgctcctagcactcctggaacagacagatcctt 978
Db 505 SYS.S.S.SMSYSYS.SSDY.CYCCYRYHCSDSYSY.Y.CRCCTYT.SYSRYDCHYSC 446
OY 979 ctggcacttctccccgggggagcacttctgtcaggaagcagcagcctcagcactctg 1038
Db 505 SYS.S.S.SMSYSYS.SSDY.CYCCYRYHCSDSYSY.Y.CRCCTYT.SYSRYDCHYSC 446
```

```
Db 445 CCSDYCYCYSYRYSYSYSYSYSYSYDYSYRCCCYRCCYVYSSSSSSSSAYST 386
OY 1039 ggcagcaacagagtccttctgacacttcccccaggaagtcacttctgtcagggagcag 1098
Db 385 SSSSSSSSSYYTSTNC.T.CC...T.MCAABCSYTTTTTTTTT.HSCC.SA..A. 326
OY 1099 ccaagcctagcagctcctgggagcaacagagtccttctggcagcccttcccaaggagca 1158
Db 325 M.YC.A.SYSYS.S.S.S.SYMR.HRA.SHYTRS..S.MCY.YM.Y.Y.Y.YSY 266
OY 1159 cttgtgctagcagatcagatgcttgcgaacatccggagtccttgcgaagcagagcagt 1218
Db 265 YCSRKTM...TMTDM.T.T..MHMY.KYB.HCHTKCRAAT.MN.HTB...N..HBHB.B 206
OY 1219 ctatgtagcagtgctgtgccccctcagattccaacccctgtcatgcccagcatcagaca 1278
Db 205 H.H..B.H.HSNS...TTS...M.TTM.B.TBASH.RARC.STMM.T.MMM.H.S.Y. 146
OY 1279 ttagtgaatgggtatccttctctccgcatacagcccaacccctgagcactggagt 1338
Db 145 .CCR..H.R..R.G...S.CT.HN.B.CYR.RNGMY.HS...S.AA...AHH..DH. 86
OY 1339 cctgtgtgacatacagaagcgg 1362
Db 85 DTYBH.YH.KNNAHC.W.KKNB.S 62
```

```
RESULT 6
US-10-141-761-10/c
; Sequence 10, Application US/10141761
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Audrey E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C198
; CURRENT APPLICATION NUMBER: US/10/141,761
; PRIOR APPLICATION REMOVED - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-141-761-10
```

```
Query Match 2.28; Score 40.6; DB 7; Length 594;
Best Local Similarity 6.98; Pred. No. 0.14;
Matches 35; Conservative 183; Mismatches 286; Indels 0; Gaps 0;
```

```
OY 859 acaatggaacatccggtatctctgacagcaacagatccttctggcacttcccccag 918
Db 565 A.MCTT..AMM.M.CBT.STT.YA.M.YT.S.S.S.SYSYS.S.S.S.SDSYSYA.SY 506
OY 919 gaggcacttctgtcaggaagcagcatgctcctagcactcctggaacagacagatcctt 978
Db 505 SYS.S.S.SMSYSYS.SSDY.CYCCYRYHCSDSYSY.Y.CRCCTYT.SYSRYDCHYSC 446
```

QY	979	cttgacactctccccaaggagctactcttctgacagggaagggcaggcagctagctactctg	1038
Db	445	CCSYYCTSYSTSYSTSYSTSYSTSWTSYSTTDCSTRRCCCTYYSSSTSSSYSSAAT	386
QY	1039	gacgaacaaaggctctctctctgcaactctccccaaggactcatttgtagcaggagca	1098
Db	385	SSSSSSSSSYTSTNMC.T.CC...T.MCAABSTTTTTTTTT.HSCC.SA..A.	326
QY	1099	ccacgcctagcaactcctctggcagaacaaagctcctctctgcaaccctctccagaagcca	1158
Db	325	M..YC.A.SYSYSYS..SSS.S.SYMR.HRA.SHYYTRS..S.MCY.YM.Y..Y.YYSY	266
QY	1159	cttgtagcagatcaagaatgatactccctcggaacatcgcgagctctctggaacgaaggacgt	1218
Db	265	YCSKRTM...TMTDM.T.T..MMH.KTB.HCHTKCPAT.MN.HTB...N..HBHB.B	206
QY	1219	ctatggacacagctgtgtgcccctcgaatccaaccctgatacgccatccatgaacaa	1278
Db	205	H.H..B.H.HSNS...TTS...M.TTM.B.TBASH.RARC.STMM.T.MMM.H.S.Y.	146
QY	1279	ttagtggaaatggtatcctctctctcaaccgatacagaagcccaaacctcgatcattggat	1338
Db	145	.CCR..H.R..R.G...S.CT.HN.B.CYRY.RNGMY.HS...S.AA...ARH..DH.	86
QY	1339	ccgtgtgtgacatacagaagaagc	1362
Db	85	DTYBH..YH.KNNAH.C.W.KNB..S.62	

```

RESULT 7
US-10-140-864-10/c
: Sequence 10. Application US/10140864
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: DeForge, Laura
: APPLICANT: Desnoyers, Luc
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerltsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Collin K
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: TITLE OF INVENTION: ACIDS ENCODING THE SAME
: FILE REFERENCE: P9330R1C184
: CURRENT APPLICATION NUMBER: US/10/140,864
: PRIORITY FILING DATE: 2002-05-07
: Prior Application removed - See Palm or File Wrapper
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO 10
: LENGTH: 594
: TYPE: PR1
: ORGANISM: Homo Sapien
: US-10-140-864-10

```

Query Match 2.2%; Score 40.6; DB 7; Length 594;
Best Local Similarity 6.9%; Pred. No. 0.14;
Matches 35; Conservative 183; Mismatches 286; Indels 0; Gaps 0;

Oy 859 acatagtgaacatccggtattctccttgagcaagaacagatccttctcgcatctttcccacg 918
| : | : | : : : : : : : : : : : : : :
Db 565 A.MCTT..AMM.M.CBT.STT.YA.M.YT.S.S.S.SYSYSYS.S.S.SDSYSA.SY 506

QY 919 gaggcactctgtgtcaaggagcagccatcgcactcctctggaagaacagatcctt 978
 Db 505 SY.S.S.WSSYSTSSSDY.CCCTCYRMCBSYSTSYT.CRCCYTT.SYSRDCHASC 446
 QY 979 ctggcactctcccccaggggtactcttgtgtcaaggaagccacgcgtacgactctg 1038
 Db 445 CCSDDYCYSYSSRYSYSYSSSYSTYDCCSYRCCCYSSSSSYSSSSAST 386
 QY 1039 ggcgcgaacaggtctcttcctcgggacctctcccccaggaatcacttgtgtcaaggacag 1098
 Db 385 SSSSSSSSYTYSTNVC.T.CC...T.MCABCSYTYTTPYTP...HSCC.SA.A. 326
 QY 1099 ccaagcctagcactcctcgtggcagaacagctcttctgtgacacctctcccaagagcca 1158
 Db 325 M..YC.A.SYSYSS.SSS.S.SYMR.HRA.SHYTTRS..S.MYCY.YM.Y..Y.YYSY 266
 QY 1159 cttgtgtcagatcaagatgactcgtggaacatccggatctcttgacagcaggaacagt 1218
 Db 265 YCSNRTM...TWTDM.T.T..MHMY.KYB.HCHTKCRAT.MN.HTB...N..HBHB.B 206
 QY 1219 ctatgtgacagtggtgtgccccccagatcccaaaccccttcatgcgcgcacatcatgaaca 1278
 Db 205 H.H..B.H.HSNS...TTS...M.TTM.B.TBASH.RARC.STMM.T.MMM.H.S.Y. 146
 QY 1279 ttaatggaatgggtactcctctcttaccgcgcatacagcccaacctcgatcatcggagt 1338
 Db 145 .CCR.H.R..R.G...S.CT.HN.B.CYRY.RNGMY.HS...S.AA...ARH.DH. 86
 QY 1339 cctgtgtgacatcacagaagacag 1362
 Db 85 DTYBH.YH.KNNAH.C.W.KNB.S. 62

```

RESULT      8
US-10-140-923-10/c
; Sequence 10. Application US/10140923
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C188
; CURRENT APPLICATION NUMBER: US/10/140,923
; PRIOR APPLICATION FILING DATE: 2002-05-07
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PRF
; ORGANISM: Homo Sapien
US-10-140-923-10

```

Query Match	2.28;	Score 40.6;	DB 7;	Length 594;
-------------	-------	-------------	-------	-------------

Desc Local Similarity: 0.98; Freq: NO: 0.14;
 Matches: 35; Conservative: 183; Mismatches: 286; Indels: 0; Gaps: 0;

QY 859 acaatggaacatccgtattccttgacagcaacagatccttctgtgcactcttccccaag 918

```

Db      565 A.MCTT...AMM.M.CBN.SWT.YA.M.YT.S.S.S.SYSYVSYS.S.S.SDSYSYA.SY 506
Qy      919 gagccacttctgttcaggacgagcgcattcgcttagcacctcttgagcaagaacagatcctt 978
Db      505 SYS.S.SMSYSYSSSDDY.CYCCKYRHGSDSYSYSYY.CRCCYRT.SYSRYDCHWSC 446
Qy      979 ctggcactctccccgggggactttgttcaggagcagcgaagcctagcactcty 1038
Db      445 CCSDIYKYSYSKNYYSYSISWMSYSTYDTCSTRRCCYYSISSSYSSSAYST 386
Qy      1039 ggcgcacaacgctctctctctgacactctcccacaggatcacttctgttcaggacag 1098
Db      385 SSSSSSSSYVNSTNYC.T.CC....T..MCABCSSTTTTTTTTTTTT.HSCC.SA.A. 326
Qy      1099 ccaagcctagcactctccgggcagacaacaggtctctcttgacacctctccaggagcca 1158
Db      325 M..YC.A.SYSYS.S.SS.S.SYMR.HRA.SHYYRS..S.MCY.YM.Y.YY.YYSX 266
Qy      1159 cttgttcaggatacgaagtatgcctggaaacatccgaggtccttgagcagcaggagagt 1218
Db      265 YCSKRMT...TWMDM.T.T..MHMY.KTB.HCHTKCRAT.MN.HTB...N..HBHB.B 206
Qy      1219 ctatggacagctgttgttgcgccctcagattccaaccctgtcatgcgcgatcatgaaca 1278
Db      205 H.H..B.H.HSNS...TTS.....M.TTM.B.TBASH.RARC.STMM.T.MM.H.S.Y. 146
Qy      1279 ttatggaaatggatcctctctcttcacgcgcatacgaagcccaacctcgatcacggagt 1338
Db      145 .CCR.H.R.R.G...S.CT.NB.CYRX.RNGMY.HS...S.AA...ARR.DH. 86
Qy      1339 cctgtgttcacatcagaagacag 1362
Db      85 DTYBH.YH.KNNAH.C.W.KNB.S 62

RESULT    9
US-10-141-756-10/c
; Sequence 10, Application US/10141756
; GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Geriltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gunney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Thomas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P333ORIC200
CURRENT APPLICATION NUMBER: US/10/141,756
Prior Filing DATE: 2002-05-08
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 10
LENGTH: 594
TYPE: PRF
ORGANISM: Homo Sapien
US-10-141-756-10
```

Query Match	2.28; Score 40.6; DB 7; Length 594;
Best Local Similarity	6.98; Pred. No. 0.14;

	Matches	35, Conservative	183, Mismatches	286, Indels	0, Gaps	0,
QY	835	acaatgagaacatcgcggtatctcctgagacagacaagatccttcctggaactcttccccag	918			
Db	565	A.MCCTT..AMM.M.CBT..STT.YA.M.YT.S.S.S.SYSYSYS.S.S.SDSYSYA..SY	506			
QY	919	ggagcaacttgtgtcaaggacagccacgactcagcactctctgtacagacagacgactctt	978			
Db	505	SYS.S.SWSYSYSSSDDI.CYCCYRRHCHSDSYSYSIYY..CRCCYTT..SYSRYDCHSYC	446			
QY	979	ctggcactcttcccccaagggtgctacttgtgtcaggagacagcagcagctagcactctg	1038			
Db	445	CCSDYCYYSYSRKRYSYSYSYSYSYSYTYDYCYRRCCTYSYSSTSSYSATYT	386			
QY	1039	ggcagcaacaaggctccttcttggaactcttccccaggaagctacttgtgtcaggaggag	1098			
Db	385	SSSSSSSSSYTSTNYC.T.CC....T..MCABCSITTTTITTTT..HSCC.SA..A..	326			
QY	1099	ccaagcctcagcactcctctgggacagacaagtgctcttctctggaaccccttccagaggca	1158			
Db	325	M..YC.A.SYSYSYS..SSS.S.SYMR.HMA..SHYTYRS..S.MCYX.YM.Y..Y.YYSY	266			
QY	1159	cttctgttcaggatacaaggtatcctcgtgacacatcggagctcgtgacagcagggacagt	1218			
Db	265	YCSKRTM...TMTDM.T.T..MMHY.KTB..HCHTKCRAT..MN..HTB..N..HBHB..B	206			
QY	1219	ctagtgagacagtgctgtgtccctcagatccaacacctgtcatcggccatccatgaaca	1278			
Db	205	H.H..B.B.H.HSNS...TTS....M.TTM.B.TBASH.RARC.SIMM.T.MMM.H.S.Y..	146			
QY	1279	ttagtggaaatgggtatcctcttctcaccgacataagcccaaacctcgatcaactgggat	1338			
Db	145	.CCR..H.R.R.G...S.CT.HN.B.CYRX.RNGMY..HS...S.AA...ARH..DH..	86			
QY	1339	cctgtgtgacatcacagaagcagg	1362			
Db	85	DYTBH..YH..KNNAH.C.W..KKNB..S..62				

```

RESULT 10
US-10-141-759-10/C
/ Sequence 10, Application US/10141759
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Beresini, Maureen
/ APPLICANT: Deforge, Laura
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tunas, Daniel
/ APPLICANT: Watanabe, Colin K
/ APPLICANT: Wood, William
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ TITLE OF INVENTION: ACIDS ENCODING THE SAME
/ FILE REFERENCE: P3330R1C197
/ CURRENT APPLICATION NUMBER: US/10/141.759
/ PRIOR FILING DATE: 2002-05-08
/ Prior Application removed - See file Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 10
/ LENGTH: 594
/ TYPE: PRT
/ ORGANISM: Homo Saplen
US-10-141-759-10

```



```
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PR
; ORGANISM: Homo Sapien
US-10-140-805-10
```

```
Query Match      2.2%; Score 40.6; DB 7; Length 594;
Best Local Similarity 6.9%; Pred. No. 0.14;
Matches 35; Conservative 183; Mismatches 286; Indels 0; Gaps 0;
```

```
Oy 859 acaatggaacatccggtatctcttgagacagcaacatctcttgacactcttcccccag 918
Db 565 A.MC.TT..AMM.M.CBT.STT.YA.M.YT.S.S.S.SYSYSYS.S.S.SDSYSA.SY 506
Oy 919 gagccacttgytcaaggacagccatgcttagacactccggcagacagacagactctt 978
Db 505 SYS.S.SMSYSYSSSDDY.CYCCYRHCSDSYSYTY.CRCCTY.SYSRYDCHYSC 446
Oy 979 ctggacactctcccccagggtcacttgytgcaggagacagccagcctagcactctg 1038
Db 445 CCSDYCYYSYSRYSYSYSYSYSYSYTDYCSYRCCCYYSYSSSYSSSAYST 386
Oy 1039 ggcagcaacagtgctctcttgacactcttcccccaggagactcttgytgcaggag 1098
Db 385 SSSSSSSSYTSTNYC.T.CC...T.MCAABCSTTTTITTTT..HSCC.SA..A. 326
Oy 1099 ccagccttagacactccgtggcagacagagctctcttgacaccccttccagagca 1158
Db 325 M.YC.A.SYSYS.S.SS.S.SYMR.HRA.SHYTRS..S.MCY.YM.Y..Y.YYSY 266
Oy 1159 ctgtgctcaggaataaggtatgctctggaacatccggaggtccttgagacagagga 1218
Db 265 YCSKRTM...TMTDM.T.T..MHMY.KYB.HCHTKCRAAT.MN.HTB...N..HBHB.B 206
Oy 1219 ctatggaacagtgtgtgcccctcagatccaaacccgtcactgcgcgcacatgaaca 1278
Db 205 H.H..B.H.HSNS...TTS...M.TTM.B.TBASH.RARC.STTM.T.MMM.H.S.Y. 146
Oy 1279 ttagtgaataggtatctctcttctacgcatacagcccaacccctgcatcagatgg 1338
Db 145 .CCR..H.R.R.G...S.CT.HN.B.CYRY.RNGMY.HS...S.AA...ARH.DH. 86
Oy 1339 cctgtgtgacatacagaagcgg 1362
Db 85 DTYBH.YH.KNNAHC.W.KKND.S 62
```

```
RESULT 13
US-10-142-885-10/c
; Sequence 10, Application US/10142885
```

```
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P33081C248
; CURRENT APPLICATION NUMBER: US/10/142,885
```

```
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PR
; ORGANISM: Homo Sapien
US-10-142-885-10
```

```
Query Match      2.2%; Score 40.6; DB 7; Length 594;
Best Local Similarity 6.9%; Pred. No. 0.14;
Matches 35; Conservative 183; Mismatches 286; Indels 0; Gaps 0;
```

```
Oy 859 acaatggaacatccggtatctcttgagacagcaacatctcttgacactcttcccccag 918
Db 565 A.MC.TT..AMM.M.CBT.STT.YA.M.YT.S.S.S.SYSYSYS.S.S.SDSYSA.SY 506
Oy 919 gagccacttgytcaaggacagccatgcttagacactccggcagacagacagactctt 978
Db 505 SYS.S.SMSYSYSSSDDY.CYCCYRHCSDSYSYTY.CRCCTY.SYSRYDCHYSC 446
Oy 979 ctggacactctcccccagggtcacttgytgcaggagacagccagcctagcactctg 1038
Db 445 CCSDYCYYSYSRYSYSYSYSYSYSYTDYCSYRCCCYYSYSSSYSSSAYST 386
Oy 1039 ggcagcaacagtgctctcttgacactcttcccccaggagactcttgytgcaggag 1098
Db 385 SSSSSSSSYTSTNYC.T.CC...T.MCAABCSTTTTITTTT..HSCC.SA..A. 326
Oy 1099 ccagccttagacactccgtggcagacagagctctcttgacaccccttccagagca 1158
Db 325 M.YC.A.SYSYS.S.SS.S.SYMR.HRA.SHYTRS..S.MCY.YM.Y..Y.YYSY 266
Oy 1159 ctgtgctcaggaataaggtatgctctggaacatccggaggtccttgagacagagga 1218
Db 265 YCSKRTM...TMTDM.T.T..MHMY.KYB.HCHTKCRAAT.MN.HTB...N..HBHB.B 206
Oy 1219 ctatggaacagtgtgtgcccctcagatccaaacccgtcactgcgcgcacatgaaca 1278
Db 205 H.H..B.H.HSNS...TTS...M.TTM.B.TBASH.RARC.STTM.T.MMM.H.S.Y. 146
Oy 1279 ttagtgaataggtatctctcttctacgcatacagcccaacccctgcatcagatgg 1338
Db 145 .CCR..H.R.R.G...S.CT.HN.B.CYRY.RNGMY.HS...S.AA...ARH.DH. 86
Oy 1339 cctgtgtgacatacagaagcgg 1362
Db 85 DTYBH.YH.KNNAHC.W.KKND.S 62
```

```
RESULT 14
US-10-146-731-10/c
; Sequence 10, Application US/10146731
```

```
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
```

```

; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C323
; CURRENT APPLICATION NUMBER: US/10/146,731
; CURRENT FILING DATE: 2002-05-15
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-146-731-10

```

```

Query Match      2.2%; Score 40.6; DB 7; Length 594;
Best Local Similarity 6.9%; Pred. NO. 0.14;
Matches 35; Conservative 183; Mismatches 286; Indels 0; Gaps 0;

```

```

QY 859 acaatgaaatccggtatctctgacgaagaatctcttctgacactctcccccag 918
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 565 A.MCTT.AMM.M.CBT.STT.YA.M.YT.S.S.S.SYSYSYS.S.S.SDSYSTA.SY 506

QY 919 gagccattgtgcaggagcagcatgctacgactcctgacagcaacagatcctt 978
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 505 SYS.S.SMSYSYSSSDY.CYCCYRHCSDSYSYTY.CRCCTYT.SYRNDCHYSC 446

QY 979 ctggacactctcccccaggagtcacttgytgcaggagcagcagcctagacactctg 1038
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 445 CCSDYCYSYSTRYSYSYSMSYSYTDYCSYRCCCYSSSYSSSYSSSYSSSYST 386

QY 1039 ggcagcaacagtccttctgacactctcccccagagtcacttgytgcaggagcag 1098
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 385 SSSSSSSSYTYSTMYC.T.CC...T.MCAABCSITTTTTTTTTT..HSCC.SA..A. 326

QY 1099 ccacgcctagcactccttctgacagcagtccttctgacactcttcccccagagca 1158
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 325 M..YC.A.SYSYSYS.SSS.S.SYMR.HRA.SHYTRS..S.MCY.YM.Y.Y.YYSY 266

QY 1159 ctgtgtcaggaatcagatgctctgacacatccgagtcctgacagcagggacagt 1218
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 265 YCSRKTW...TWTDW.T.T.MMYT.KYB.HCHTRCRAAT.MN.HTB...N..HBHB.B 206

QY 1219 ctatgtgacagtgctgtgcctccctcagatcccaacccctgcatgcgcacatcattga 1278
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 205 H.H..B.H.HSNS...TTS....M.TTM.B.TBASH.RARC.STMM.T.MMM.H.S.Y. 146

QY 1279 ttatgtgaaatgggtatcctctctcaccgcatacagcccaacccctcgatcagat 1338
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 145 .CCR..H.R..R.G...S.CT.HN.B.CYRY.RNGMY.HS...S.AA...ARH..DH. 86

QY 1339 cctgtgtgacatcacagaagcag 1362
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 85 DTYBH.YH.KNNAHC.W.KKNB.S.62

```

```

RESULT 15
US-10-142-426-10/c
; Sequence 10, Application US/10142426
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K

```

```

; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C224
; CURRENT APPLICATION NUMBER: US/10/142,426
; CURRENT FILING DATE: 2002-05-09
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-142-426-10

```

```

Query Match      2.2%; Score 40.6; DB 7; Length 594;
Best Local Similarity 6.9%; Pred. NO. 0.14;
Matches 35; Conservative 183; Mismatches 286; Indels 0; Gaps 0;

```

```

QY 859 acaatgaaatccggtatctctgacgaagaatctcttctgacactctcccccag 918
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 565 A.MCTT.AMM.M.CBT.STT.YA.M.YT.S.S.S.SYSYSYS.S.S.SDSYSTA.SY 506

QY 919 gagccattgtgcaggagcagcatgctacgactcctgacagcaacagatcctt 978
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 505 SYS.S.SMSYSYSSSDY.CYCCYRHCSDSYSYTY.CRCCTYT.SYRNDCHYSC 446

QY 979 ctggacactctcccccaggagtcacttgytgcaggagcagcagcctagacactctg 1038
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 445 CCSDYCYSYSTRYSYSYSMSYSYTDYCSYRCCCYSSSYSSSYSSSYSSSYST 386

QY 1039 ggcagcaacagtccttctgacactctcccccagagtcacttgytgcaggagcag 1098
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 385 SSSSSSSSYTYSTMYC.T.CC...T.MCAABCSITTTTTTTTTT..HSCC.SA..A. 326

QY 1099 ccacgcctagcactccttctgacagcagtccttctgacactcttcccccagagca 1158
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 325 M..YC.A.SYSYSYS.SSS.S.SYMR.HRA.SHYTRS..S.MCY.YM.Y.Y.YYSY 266

QY 1159 ctgtgtcaggaatcagatgctctgacacatccgagtcctgacagcagggacagt 1218
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 265 YCSRKTW...TWTDW.T.T.MMYT.KYB.HCHTRCRAAT.MN.HTB...N..HBHB.B 206

QY 1219 ctatgtgacagtgctgtgcctccctcagatcccaacccctgcatgcgcacatcattga 1278
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 205 H.H..B.H.HSNS...TTS....M.TTM.B.TBASH.RARC.STMM.T.MMM.H.S.Y. 146

QY 1279 ttatgtgaaatgggtatcctctctcaccgcatacagcccaacccctcgatcagat 1338
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 145 .CCR..H.R..R.G...S.CT.HN.B.CYRY.RNGMY.HS...S.AA...ARH..DH. 86

QY 1339 cctgtgtgacatcacagaagcag 1362
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 85 DTYBH.YH.KNNAHC.W.KKNB.S.62

```

Search completed: September 1, 2002, 02:10:33
Job time: 6037 sec